ELSEVIER

Journal of Chromatography B, 782 (2002) 267–289

JOURNAL OF CHROMATOGRAPHY B

www.elsevier.com/locate/chromb

Multidimensional chromatography coupled to electrospray ionization time-of-flight mass spectrometry as an alternative to two-dimensional gels for the identification and analysis of complex mixtures of intact proteins

Hongji Liu, Scott J. Berger, Asish B. Chakraborty, Robert S. Plumb, Steven A. Cohen*

Life Sciences R&D Group, Waters Corporation, 34 Maple Street, Mail Stop TG, Milford, MA 01757, USA

Abstract

The limitations of 2-D gels for global proteomics have encouraged the development of alternative approaches for identifying proteins in complicated mixtures, and determining their modification state. In this work, we describe the application of multidimensional liquid chromatography (SCX-RPLC) coupled with electrospray time-of-flight mass spectrometry and off-line fraction collection to analyze complex intact protein mixtures. Methods were developed using both standard proteins and an enriched yeast ribosomal fraction sample containing ~100 proteins, which permitted assessment of the effectiveness of the individual separation dimensions, as well as investigation of the interplay between separation capacity and electrospray MS performance.

Keywords: Proteomics; Ribosomes; Proteins

1. Introduction

In both global and targeted proteomic studies, the limitations of 2-D gel electrophoretic methods are stimulating the development of non-gel based alternatives. Recent reports from several groups have demonstrated the utility of 1-D and 2-D chromatographic methods for the separation of global protein digests [1–4], but chromatographic methods for complex mixtures of intact proteins are applied far less frequently to proteomic studies. Recent developments in global protein separations for proteomics include reports on simplifying whole cell extracts prior to 1- or 2-D gels [5], and application of on-line electrophoretic [6,7] and chromatographic separations [8,9] coupled to mass detection.

Whole cell extracts contain a diverse set of proteins, which vary from hydrophilic to hydrophobic (including integral membrane proteins), highly acidic to neutral to highly basic, and with predicted masses ranging from below 10 kDa to in excess of 150 kDa. This diversity of molecular properties results in a level of complexity where a single mode of chromatography produces fractions enriched with similar components, but where the sheer number of proteins and modified proteins prevents efficient detection and characterization of individual components. To date, no single dimension

*Corresponding author. Tel.: +1-508-482-2501; fax: +1-508-482-3625.

E-mail address: steven_cohen@waters.com (S.A. Cohen).

1570-0232/02/$ – see front matter
PII: S1570-0232(02)00554-8

HPLC separation method can remotely approximate the effective resolving power of 2-D gel systems in routine use today [10].

Multi-dimensional HPLC separations and coupled HPLC–CE systems are currently being investigated as promising alternatives for intact protein separations in proteomics. The basic theory behind multi-dimensional separations, described by Giddings in the 1980s [11,12], and expanded upon by others [13,14], shows that significant increases in system peak capacity occur with 2-D systems having orthogonal separation mechanisms. Jorgenson's group has described several 2-D LC systems for protein separations. In one system combining IEX with RPLC, the effluent from the initial IEX step alternately delivered to one of two loading loops connected to a valve located between the IEX and RPLC columns [15]. After the filling the first loop with IEX effluent, the valve was used to direct the RPLC gradient though the loop onto the RP column, placing the second loop in-line with the IEX column. The process continued until the end of the IEX gradient. In a second system coupling SEC with RPLC, a valve-switching method was employed to divert the effluent from the first separation dimension to two alternating RPLC columns for the second dimension [10,16]. The SEC–RPLC system was also interfaced with ESI-MS detection for on-line identification of protein mixture components.

The most intensive studies to date have been reported by Lubman and colleagues [8,9,17] who have coupled a chromatofocusing first dimension with a reversed-phase LC second dimension for the characterization of proteins in cancer cell extracts. In this work, protein masses were determined on-line by ESI–TOF MS, while peptide analysis of digested split fractions confirmed identifications. More recently Unger and co-workers [18–20] described two- and three-dimensional chromatographic systems that combine size separation, ion-exchange and reversed-phase modes to fractionate peptides and small proteins from partially purified hemofiltrate. In this case, fractions collected from the final reversed-phase step were analyzed using MALDI–TOF MS.

In this report, we have focused on optimizing a comprehensive two-dimensional HPLC system for intact proteins incorporating an initial ion-exchange step and a second dimension reversed-phase separation, which permits both fraction collection and on-line mass analysis using an ESI–TOF MS system. Using columns suitable for the analysis of a wide range of proteins, we demonstrate the system's utility to resolve mixtures of standard proteins, and more complex protein mixtures present in an enriched yeast ribosomal protein fraction. Comparisons of the 2-D system with a single dimension RPLC separation conclusively demonstrate the power of the orthogonal IEX–RPLC system to present less complex fractions to an analytical detector, with a resultant S/N enhancement for individual sample components. The volatile mobile phase used for the final reversed-phase step permitted direct coupling to ESI–TOF MS detection, which yielded tentative assignments for closely related molecular species, including modified proteins, and nearly identical protein isoforms.

2. Experimental

2.1. Chemicals

Acetonitrile was purchased from J.T. Baker (Pittsburgh, USA), while formic acid, trifluoroacetic acid (TFA), standard proteins, and buffer components were obtained from Sigma–Aldrich (St. Louis, MO). Class I (18 MΩ) deionized water was produced in house using an ElixS/GradientA10 (Millipore) water purification system.

2.2. Ion-exchange chromatography

LC separations were performed using one or two Waters 2795 liquid chromatographs (Waters, Milford, MA), where eluted components were detected by UV absorbance at 280 nm with a Model 2487 Dual Wavelength Detector (Waters) and component masses determined using an LCT™ ESI–TOF MS (Micromass UK). Cation-exchange chromatography was performed on a Shodex SP-420N column (4.6× 35 mm; Showa Denko, Japan), and components were eluted with gradient profiles as described below. Background contaminants originating from the urea used in the ion-exchange buffers were removed by passing mobile phases once through a 5 cm×7.8 mm Symmetry® C_{18} (Waters) column.

2.3. Reversed-phase chromatography

Reversed-phase chromatography was performed on either an 4.6×33 mm Eichrom NPS C_{18} 1.5 μm (Darien, IL, USA) or a 2.1×50 mm Symmetry300™ C_4 column (Waters). Columns maintained at room temperature were developed using gradient conditions of 25–60% B at a flow-rate of 0.5 ml/min or 10–60% B at a flow-rate of 0.5 ml/min, where eluent A was 0.1% TFA in water and eluent B was 0.1% TFA in acetonitrile. The gradient time was varied from 10 to 200 min and the column effluent was monitored by UV at 280 nm.

2.4. Two-dimensional chromatography

Two-dimensional chromatography (SCX-RPLC) was operated in two different modes. In the first mode, the ion-exchange column was developed with a linear gradient, and the effluent was alternately trapped on a pair of reversed-phase columns, which were developed using gradient elution. In this application, the reversed-phase columns were operated with 18-min gradients running from 20 to 50% B.

In the second mode of operation the ion-exchange column was developed using step elution, with each step being alternately trapped on two RP columns, which were developed using identical gradient conditions as with the continuous elution system. A schematic of the chromatographic system is shown in Fig. 1. In this mode of operation, ten step ion-exchange fractions were successively captured onto the alternate reversed-phase columns, and while a fraction was collected onto reversed-phase column 1, the second reversed-phase column was developed with the acetonitrile gradient and its flow directed to UV/fraction collection and MS detectors. A column selection valve (1) allowed the second column to be switched to receive the SCX effluent while the other reversed-phase column was developed. The system continued alternately collecting and eluting fractions until completion of the final the ion-exchange step. It is important to note that directing the reversed-phase gradient to waste during the initial 3–6 min, and final 0.5 min, of the gradient was necessary to prevent non-volatile buffer components from ion exchange from fouling the MS source, and reducing spectral quality.

2.5. Two-dimensional chromatography: model soluble proteins

A solution of bovine serum albumin (BSA) and horse heart cytochrome *c* was prepared in water to give final concentrations of 66.8 and 10.0 mg/ml, respectively. The separation of BSA and cytochrome *c* was performed using either linear or step gradient elution between 0 and 65% B with a flow-rate of 0.4 ml/min, where mobile phase A was 0.1% formic acid containing 10% acetonitrile and mobile phase B was 50 m*M* phosphate, pH 7.0 containing 1 *M* NaCl and 10% acetonitrile. Over a total run time of 84 min, the eluent was divided into fourteen 6-min fractions and collected onto two NPS C_{18} 4.6×33 mm columns, which were alternately developed with 4-min gradients (10–60% B, 1 ml/min).

2.6. Two-dimensional chromatography: enriched ribosomal protein fraction

Ribosomal proteins were separated in the first dimension of the 2-D LC system using a Shodex SP-420N 4.6×35 mm and a step gradient of 0–90% B in nine steps over 200 min at a flow-rate of 400 μl/min, using a variation of a previously published chromatographic study [21]. Eluent A was 50 m*M* methylamine, 6 *M* urea, 0.5 m*M* DTT, adjusted to pH 5.6 with acetic acid, containing 10% (v/v) acetonitrile. Eluent B was eluent A plus 1 *M* NaCl. The individual elution steps were focused alternately onto one of two identical reversed-phase columns (Waters Symmetry300 C_4, 3.5 μm, 2.1×50 mm), which were developed with a 20–50% B gradient over 18 min at 0.5 ml/min, where eluent A was 0.1% TFA in water and eluent B was 0.1% TFA in acetonitrile. The gradient tables for both the ion-exchange and reversed-phase components of the two-dimensional separation are given in Table 1.

2.7. Purification of large and small ribosomal subunits from Saccharomyces cerevisiae

Haploid yeast (*S. cerevisiae* strain BY4739 [MATα leu2Δ0 lys2Δ0 ura3Δ0] were grown to mid-log phase (OD_{600} 6.0) in shaker culture (SD media (Qbiogene), 250 rpm, 30 °C). A cell pellet was obtained by centrifugation (5000 *g*, 5 min, 4 °C) and

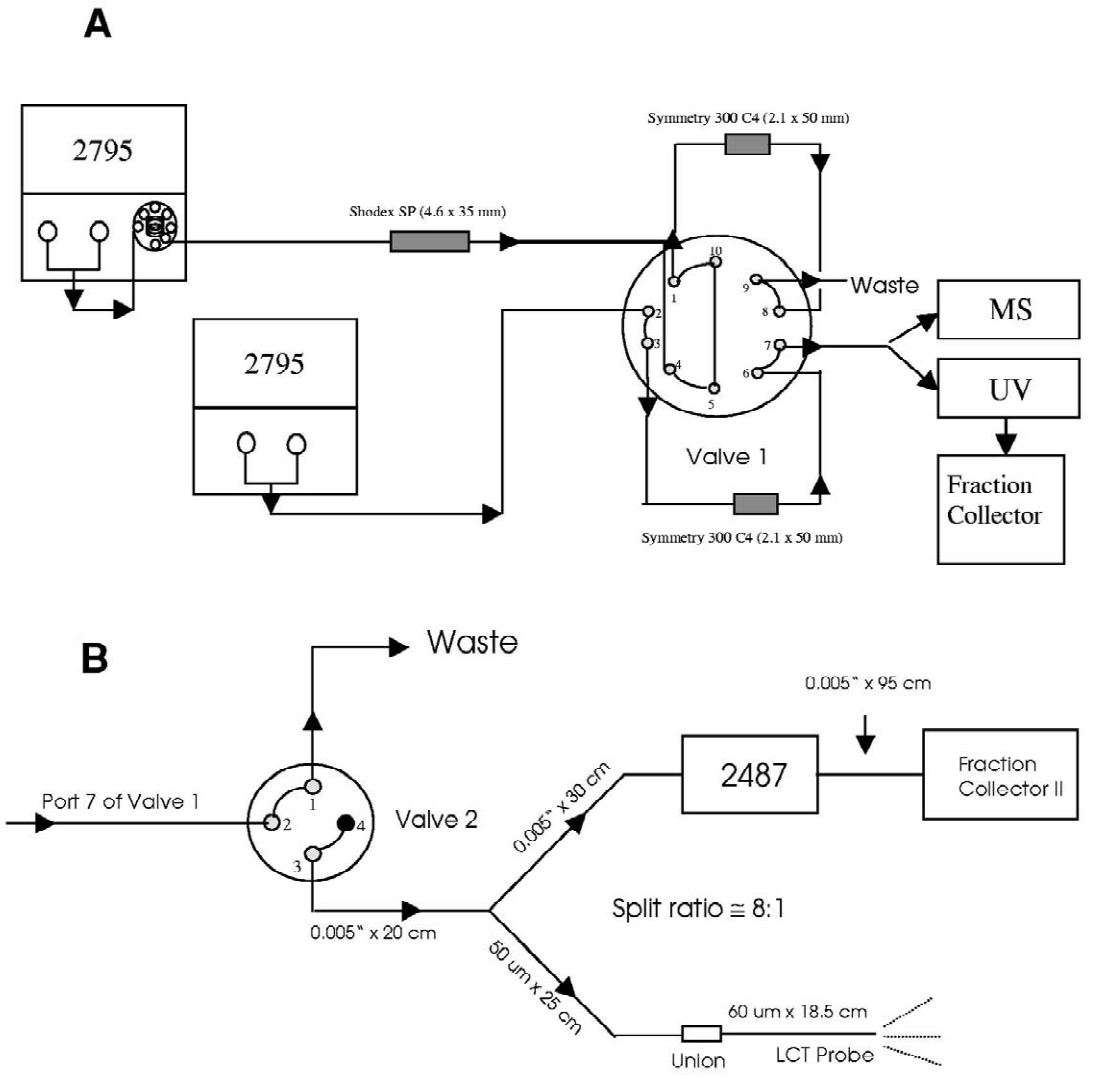

Fig. 1. Schematic of the two- dimensional separation system. (A) The complete LC–LC-UV MS system and (B) detailed view of the post-column configuration.

was washed once with cold distilled water. Ribosomal large subunits were purified using French press lysis and differential centrifugation/ultracentrifugation as described previously [22], except that ribosomes were pelleted through a salt/sucrose cushion using fixed angle centrifugation (Ti90 Rotor, 40 000 rpm, 330 min, 4 °C). In contrast to the published procedure, pelleting of ribosomes using the fixed angle rotor yields both large and small ribosomal subunits, while use of the swinging bucket rotor under essentially identical conditions (SW41, 40 000 rpm, 330 min, 4 °C) yielded only the large subunit (data not shown), as observed previously [22]. Ribosomal protein concentration (10 mg/ml) was determined using the Bio-Rad (Hercules, CA) protein assay and a bovine IgG standard. Samples were quick-frozen and stored as aliquots at −80 °C until needed.

Table 1
Gradients used for two-dimensional ribosomal protein separation

Time	Flow	% A	% B	Curve[a]
A. Ion-exchange gradient				
0	0.4	100	0	–
20	0.4	95	5	1
40	0.4	84	16	1
60	0.4	80	20	1
80	0.4	76	24	1
100	0.4	72	28	1
120	0.4	68	32	1
140	0.4	64	36	1
160	0.4	60	40	1
180	0.4	10	90	1
200	0.4	100	0	1
B. Reversed-phase gradient				
0	0.5	80	20	–
18	0.5	50	50	6
20	0.5	80	20	1
38	0.5	50	50	6
40	0.5	80	20	1
58	0.5	50	50	6
60	0.5	80	20	1
78	0.5	50	50	6
80	0.5	80	20	1
98	0.5	50	50	6
100	0.5	80	20	1
118	0.5	50	50	6
120	0.5	80	20	1
138	0.5	50	50	6
140	0.5	80	20	1
158	0.5	50	50	6
160	0.5	80	20	1
178	0.5	50	50	6
180	0.5	80	20	1
198	0.5	50	50	6
200	0.5	80	20	1

[a] Curve 1 is a step function and curve 6 is a linear gradient step.

2.8. Chromatographic sample preparation for ribosomes

Acid extraction based on the work of Hardy and coworkers [23] was performed to remove nucleic acid contaminants prior to each analysis. Ribosomal proteins were combined with 0.1 vol. of 1 *M* $MgCl_2$ and 2 vol. of glacial acetic acid, and briefly mixed several times by light vortexing over 5 min. The insoluble RNA fraction was pelleted by centrifugation (Beckman-Coulter Model 22R, 10 000 *g*, 3 min, RT), and soluble protein extracts (est. ~3.3 mg/ml protein) were stored at 4 °C prior to injection, or quick-frozen for later use.

2.9. Mass spectrometry

LC/MS analysis was performed on a LCT ESI–TOF MS equipped with an orthogonal electrospray interface (Micromass, Manchester, UK). The instrument was operated in positive ion mode with the following instrumental conditions: source temperature, 130 °C; desolvation temperature, 250 °C; capillary voltage, 3.0 kV; cone voltage, 60 V; extraction cone, 3 V; and the RF, 450 V. Nebulization gas was operated at 100 l/h while the desolvation gas was set to 400 l/h. The TOF detector acquired data for a maximum flight time of 65 μs with the MCP operated at 2700 V.

2.10. Mass spectral deconvolution and ribosomal subunit assignment

Deconvoluted masses of sample components were determined by a combination of manual and automated application of a maximum entropy (MaxEnt1) deconvolution algorithm included in the MassLynx™ ver. 3.5 software package (Micromass). MaxEnt™ spectra were generated from the entire acquired m/z range to produce deconvoluted neutral mass spectra in the mass range of 3000–46 000 Da, using a bin size of 0.5 Da, a Gaussian damage model with a width of 0.75 Da, and a processing limit of ten iterations or model convergence. LC/MS data were processed in an automated fashion by use of OpenLynx™ Software (Micromass), which automatically selected TIC peaks using the Peak ApexTrack™ function, combined spectra comprising the approximate full width of each peak, and generated deconvoluted spectra using the MaxEnt1 algorithm.

Ribosomal protein assignment was accomplished by comparison of observed masses to those predicted from the yeast genome [24] sequence assuming the presence or absence of the initiating methionine and/or *N*-acetylation. Additional identifications of processed and post-translationally modified ribosomal subunits was made based on data obtained from the Yeast Protein Database [25,26] and other sources as referenced.

2.11. Trypsin digestion

The collected ribosomal protein fractions, split from a SCX-RP-ESI–TOF MS experiment, were dried by vacuum centrifugation, and then dissolved in 50 m*M* NH_4HCO_3 containing 2 m*M* DTT and 0.1% *Rapi*Gest™ SF (Waters). Porcine trypsin (Promega, sequencing grade) was added at 1:50 (w/w) ratio and incubated for 1 h at 37 °C. Digest was stored at −80 °C until analyzed.

2.12. Peptide mass fingerprinting by MALDI–TOF MS

Peptide mass fingerprint (PMF) analyses of each of the digested fractions was performed on a Micromass M@LDI™ R TOF MS equipped with time lag focusing, and operated in the positive ion reflector mode. Acquired spectra were the sum of ten laser shots per position and ten positions per sample well. Spectra were internally calibrated using 40 fmol of [Glu1]-fibrinopeptide (Sigma) spiked into the matrix solution, as well as a major trypsin autoproteolytic peptide (MH+ = 2211.105). A 1-μl aliquot of digest was applied on the MALDI target plate followed by an equal volume of alpha-cyano-4-hydroxycinnamic acid (10 mg/ml) in 49.95% methanol, 49.95% acetone and 0.1% TFA. Automated data acquisition and processing of the PMF data were performed under Mass Lynx™ v3.5 software control. Proteins were identified by the PeptideAuto PMF database-searching module, operating within ProteinLynx Global Server™ v1.1 (Micromass). Peptide MS spectra were searched against a non-redundant translated yeast ORF database (ftp://genome-ftp.stanford.edu/pub/yeast/yeast_protein/) with constraints of 30-ppm mass measurement accuracy, potential Met oxidation, and up to one missed trypsin cleavage.

3. Results and discussion

3.1. Ion-exchange chromatography of proteins

Although both anion and cation-exchange HPLC were viable options for protein ion-exchange separations, cation-exchange HPLC was selected for the first separation dimension primarily on the basis of buffer pH compatibility with the subsequent reversed-phase dimension, and the use of positive ion mode mass detection. Buffers for optimal operation of protein anion-exchange HPLC are usually in the pH range of 7 or greater, which can cause column degradation and poor chromatographic reproducibility with silica based reversed-phase columns. An additional consideration for this choice is the presence of nucleic acids in complex biologically derived samples (e.g. cell lysates or cytosolic fractions), which would dominate competition for charged sites on SAX media, and potentially produce secondary modes of protein retention (DNA/RNA affinity, and polyphosphate based weak cation exchange).

3.2. Linear versus step gradients for ion-exchange separations

Although continuous linear gradients are the norm with protein ion-exchange chromatography, there are two strong arguments against using linear gradients and in favor of a step gradient approach. First, capturing IEX effluent on alternating reversed-phase columns simply recombines any components resolved during the ion-exchange separation, negating additional selectivity generated from application of a gradient. More importantly, ‘‘fraction splitting’’, which occurs when peaks elute as the flow path is switched from one reversed-phase column to the other, is more likely in multidimensional separations when peaks are broad. It is also directly related to the frequency of second dimension column switching such that the advantage of rapid sampling in the first dimension may be negated.

To minimize the likelihood of the fraction splitting phenomenon, we explored step gradient elution for the ion-exchange separation. This allows one to exploit the tendency of proteins to interact with ion-exchange surfaces at multiple binding sites [27,28]. In accordance with this stoichiometric displacement model is the observation that proteins will elute from an ion-exchange column at a buffer ionic strength that may be only 0.01 *M* greater than one at which retention is essentially infinite due to the cooperative nature of multiple site binding. By operating the ion-exchange column in a series of stepped increments, the theory would predict that

only those proteins with significant linear velocity at a particular ionic strength will elute in that IEX fraction, and that neighboring step fractions should be essentially devoid of those proteins.

The advantages of a step gradient versus a linear gradient are illustrated in the separation of BSA shown in Fig. 2. The separation of two BSA isoforms is greatly improved by the use of a step gradient, where both the peak shape and resolution are superior to that observed using linear gradient elution. It is worth noting that given the broad peaks observed (peak width of 2.2–5.3 min) with the linear gradient, there is a high probability of splitting a peak into two or more fractions. In contrast, each isoform is resolved into individual fractions using a step gradient mode. This observation has been made when analyzing several model proteins including cytochrome *c* and lysozyme (data not shown).

3.3. Reversed-phase separations of proteins

Theory predicts that maximum peak capacity in comprehensive 2-D separations is achieved using a sampling time for the second dimension that allows for a minimum of three fractions per peak [13]. However, there are two significant practical limitations to reducing the time of the reversed-phase step. First, there is a finite period of time to deliver the gradient to the column (system delay) as well as the

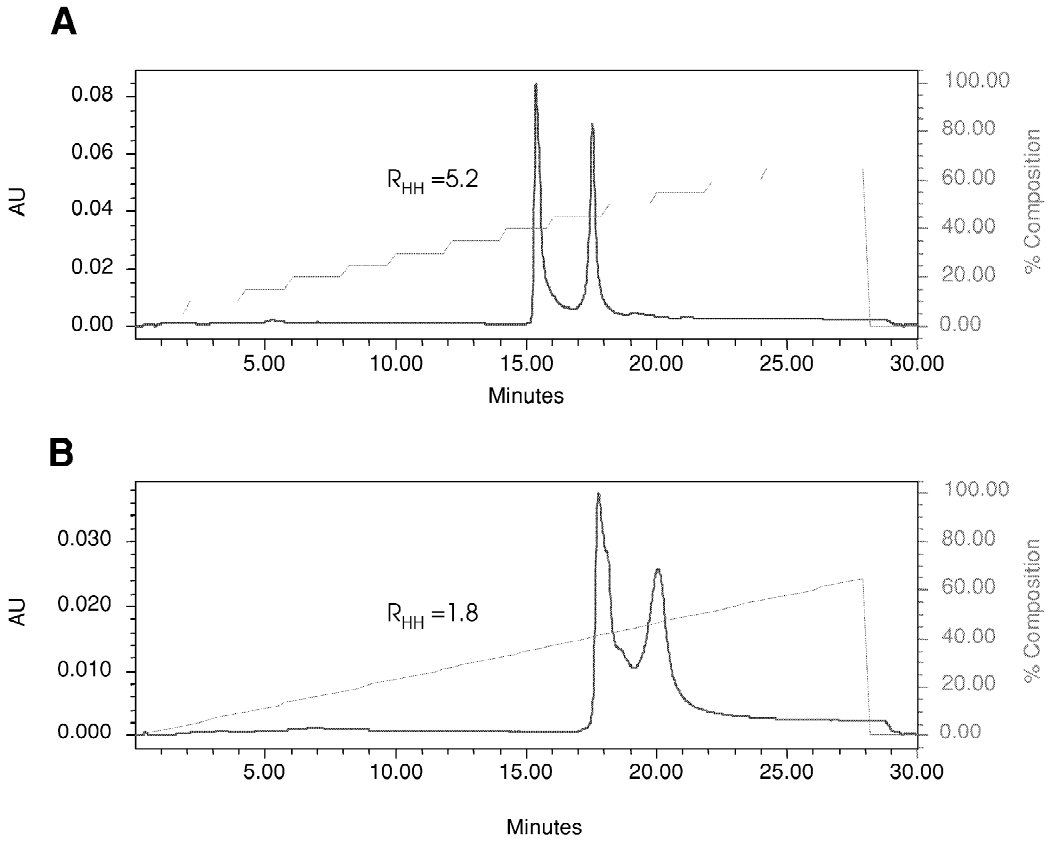

Fig. 2. Chromatogram of BSA using a step or linear ion-exchange gradient. A 2-μl injection of a BSA solution (66 mg/ml) was made onto a Shodex SP-420N column (4.6×35 mm). The column was eluted with either (A) a step gradient or (B) a linear gradient profile with eluent A=0.1% formic acid containing 10% acetonitrile and eluent B=50 m*M* phosphate, pH 7.0 containing 1 *M* NaCl and 10% acetonitrile. Separation utilized a flow rate of 1 ml/min, and UV detection at 280 nm.

need to re-equilibrate the system after each step. Furthermore, when using MS detection, the reversed-phase flow-rate is limited to a maximum of 1 ml/min to provide optimal MS sensitivity, and avoid the necessity of discarding the vast majority of the effluent to waste via a post-column splitter. Second, using the optimum sample rate would also further complicate post-run MS analysis as each protein would be present in at least three IEX fractions. This would significantly increase the post-processing burden for spectral deconvolution by several fold.

The reversed-phase gradient was optimized using a four-protein mixture (BSA, lysozyme, cytochrome *c* and carbonic anhydrase). Various gradient profiles were studied using a 2.1× 50 mm Symmetry300 C_4 column containing 5-μm particles (data not shown). These results indicated that acceptable resolution could not be obtained with short gradient times (<2 min), but that baseline resolution could be obtained with longer separations (>5 min). The necessity of scaling gradient times proportionately with column length, sets practical limits on column choice, and analysis time for the second dimension separation. Increasing the start point of the gradient to contain >10% acetonitrile, permitted shallower gradients, and increased resolution, without lengthening gradient time. The conditions for analysis of this mixture resulting in the best combination of peak resolution and cycle time was determined to be a 10–60% B gradient over 10 min or longer, with a minimum system re-equilibration time of 2 min.

Based on previous reports using reversed-phase chromatography in proteomics studies [8,9], we also compared separations using 1.5-μm non-porous silica packing with roughly comparable porous materials (data not shown) as described in Section 2.3 of the Experimental section. It is worth noting that in columns with roughly equivalent dimensions (Symmetry300 3.5 μm 4.6×50 and 4.6×53 mm 1.5 μm NPS), operating at a flow-rate of 0.5 ml/min, the NPS silica material generated a backpressure of 3400 p.s.i. compared to 200–300 p.s.i. for porous silica. This flow resistance severely limits the ability to connect the NPS columns in 2-D LC systems, where the reversed-phase column is connected in series to an ion-exchange column typically operating at flow-rates on the order of 0.4–1.0 ml/min.

3.4. Reversed-phase LC acid modifiers

The selection of the reversed-phase LC eluent is very important in LC/MS. Historically TFA has been the acidic modifier of choice for the analysis of peptides and proteins, providing good peak shape and recovery, plus the added benefit of volatility to facilitate easy removal. The recent use of ESI-MS for the analysis of proteins and peptides has resulted in this selection being re-evaluated. Several studies have indicated that formic acid significantly increases mass spectrometer sensitivity for peptide and protein analyses [29,30], while TFA has been observed to suppress MS signal to a far greater extent. In our studies (Fig. 3), we found that mobile phase containing 2% formic acid (aq) resulted in three-fold greater MS sensitivity compared to mobile phase containing 0.1% TFA (aq). This was in marked contrast to a previous report showing a 35-fold or greater increase in response with formic acid compared to TFA [29], which was observed using differing instrumentation and methods. However, we also observed that replacement of TFA with formic acid resulted in a significant reduction in chromatographic performance (Fig. 3). This is not as significant for peptides, where reversed-phase separations in low TFA, acetic acid, or formic acid result in minimal chromatographic penalty to obtain increased MS sensitivity. Our results indicate that this is not the case with proteins and where significant resolution sacrifices, and increased peak widths, would prove unacceptable for the minor gains in sensitivity, and limits on sample complexity. As discussed further below, when analyzing complex mixtures with ESI-MS detection, there is also an added benefit to high separation efficiency by eliminating competition for charging in the ESI source, and reducing instances of charge state deconvolution in the presence of multiple overlapping signals.

3.5. Peak capacity of 1-D reversed-phase separations of ribosomal proteins

Peak capacities for the 1-D and 2-D systems were determined using the equation described in Ref. [31] with an enriched fraction of yeast ribosomal proteins. The number of ribosomal components in the purified

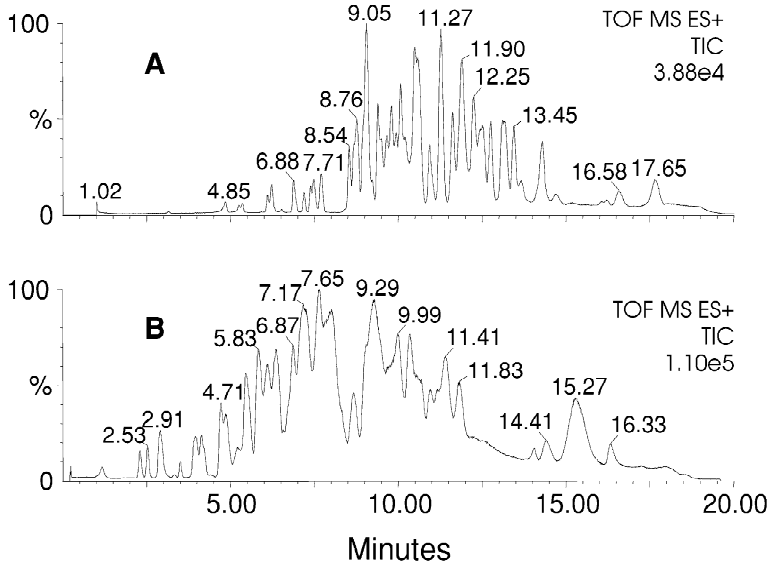

Fig. 3. Comparison of the chromatographic separation of a ribosome protein mix using either (A) 0.1% TFA or (B) 2% formic acid as the aqueous buffer. A 25-μl injection of a ribosome solution was made onto a Waters Symmetry300 C_4 3.5-μm column. The column was eluted with a 10–60% B (containing the modifier in acetonitrile) gradient over 18 min. The column eluent was monitored by positive ion ESI-MS.

yeast ribosome can vary between 78 (the number of known ''canonical'' [32,33] subunits), and 116 (the number of unique subunit isoforms) depending on isoform expression levels. The presence of multiple modified forms for some proteins [22], and the more recent identification of additional ribosomal interacting proteins such as ASC1(BEL1) [1] produce additional complexity. These proteins exhibit a size range of 3–45 kDa (mean MW~17 700), and while primarily basic (mean p*I*~10.1) the mixture potentially contains at least 14 components with predicted p*I*'s of 4.0–7.0. While complicated, this mixture of proteins represents an idealized sample in the respect that purified components are expected at roughly equimolar amounts.

In the 1-D reversed-phase system, 25 μg of ribosomal proteins were analyzed using a 10–60% B gradient at a flow-rate of 0.5 ml/min developed over 198 or 18 min (Fig. 4). As expected, resolution increases with the longer gradient; for example, the three peaks at eluting at 96.9–98.5 min on the longer gradient are completely resolved, while effectively coeluting at ~10.5 min during the shorter gradient. Using an average peak width (0.48 min) from the longer reversed-phase run, the peak capacity was determined to be ~240.

3.6. Considerations for multidimensional chromatography

The true power of applying multidimensional chromatography to complex samples is that sample components may not differ sufficiently in their relative characteristics on an individual mode of separation. The ability to increase theoretical separation peak capacity in a 1-D RPLC separation by means of longer columns, smaller particles, and shallower/longer gradients only increases separation efficiency, without actually addressing the underlying question of whether selectivity between components differs sufficiently to permit resolution. Furthermore, current success in achieving high 1-D RPLC peak capacity of biopolymers [34–36] have so far come at the cost of using pressures that limit separation component selection, separation timescales that significantly impact sample throughput, and instrumentation not readily adoptable by the larger bioanalytical community. In the following

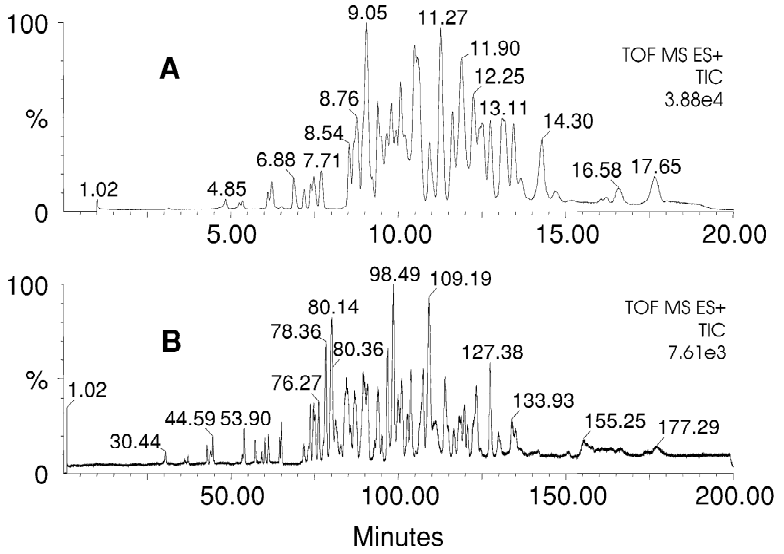

Fig. 4. Reversed-phase chromatographic separation of a ribosome protein mix using (A) an 18- and (B) a 198-min gradient. A 25-μl injection of ribosome solution was made onto a Waters Symmetry300 C_4 column that was eluted with a 10–60% B gradient, where A = 0.1% TFA in water and B = 0.1% TFA in acetonitrile. The column eluent was monitored by positive ion ESI-MS.

sections we have attempted to increase peak capacity using two orthogonal separation mechanisms (SCX and RPLC) that can be implemented using commercially available components, and under conventional pressure regimes, to evaluate peak capacity and performance relative to single dimension separations of comparable duration.

3.7. System configuration for 2-D LC/MS

Optimum performance in the 2-D LC/MS required careful attention to maintaining peak fidelity, while preventing non-volatile components entering the MS source. Controlling system bandspread is critical to maintaining peak fidelity in a system with multiple valves, fluidic connections, and post column flow splitter in a separation where peak volumes from the second dimension reversed-phase separation are ~200 μl. In our optimized system, post column tubing up to the 1:8 flow splitter was 0.005″ I.D. PEEK and the post-split to MS source connection was accomplished using a 50-μm I.D. polyimide fused-silica capillary. Appropriate attention to making proper component connections resulted in minimal post-column bandspread, as evidenced by the roughly equivalent resolution obtained by MS compared to UV detection (Fig. 5).

Also essential to obtaining good mass spectra was the diversion of non-volatile components from the ion-exchange dimension prior to directing reversed-phase column effluent into the MS source. Such non-volatile components (e.g. urea and buffer salts) were observed to result in physical blockage of the microspray/nanospray source, causing significant suppression of analyte ionization, and reduced sensitivity by generating multiple adducted peaks for each component. Although IEX effluent was directed to waste during capture of proteins onto the RP column, residual IEX eluent present in the column when it was placed in line with the MS resulted in spectra containing a high number of adduct peaks (Fig. 6). Optimizing diversion valve timing as described in the Experimental section, such that all ion-exchange buffer captured in the reversed-phase column was directed away from the MS system, minimized these adducts and enhanced sensitivity (Fig. 6, insets), while resulting in no loss of component detection. In this example, maximum relative signal (i.e. 100%) is roughly 2-fold higher for the raw mass spectra (B vs. A in Fig. 6), and more than

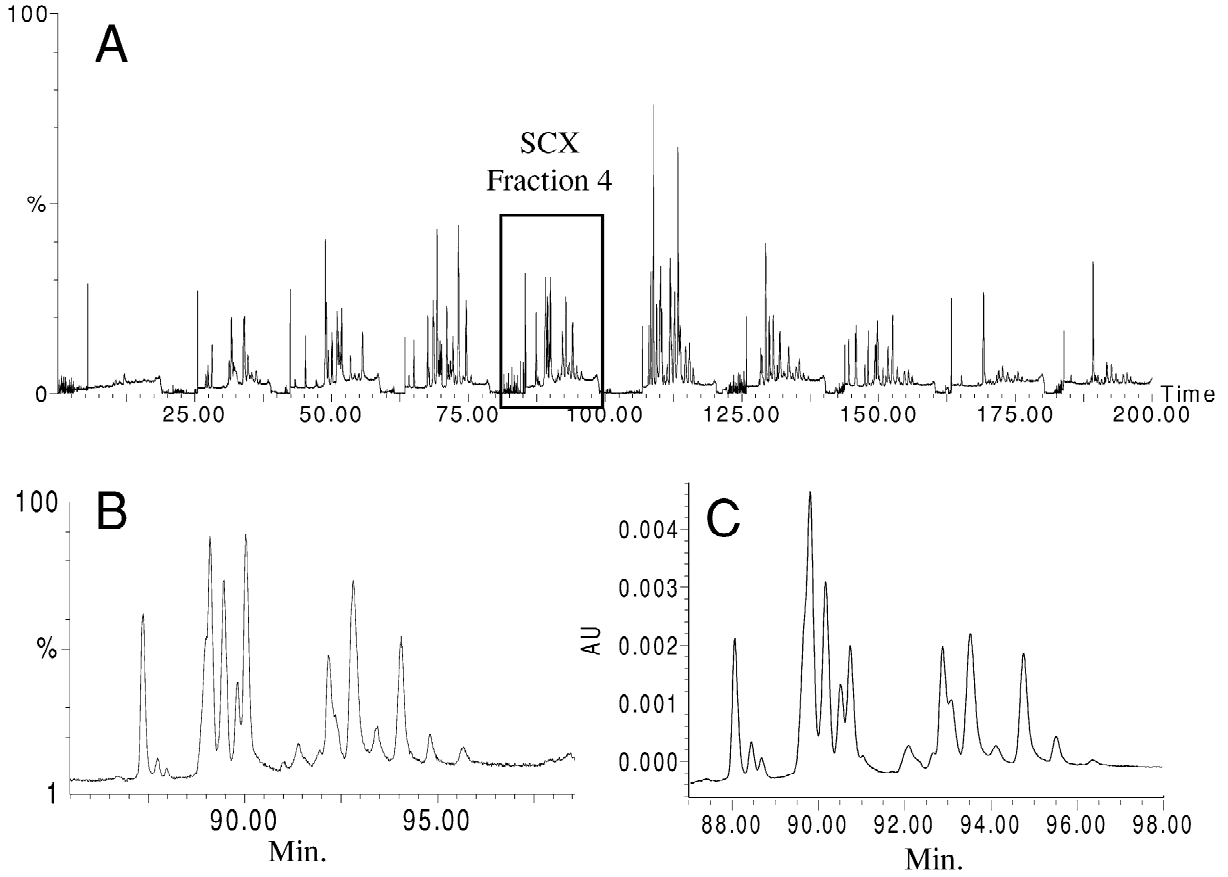

Fig. 5. Two-dimensional separation of yeast ribosomal proteins with MS and UV detection. Analysis of this sample yielded a complex TIC (panel A), where comparable peak widths were observed in the ESI–TOF MS TIC (panel B) and UV chromatogram (panel C), demonstrated using an enlarged segment of the analysis concurrent with the RP analysis of the fourth SCX step fraction.

4-fold greater for the deconvoluted spectra (B vs. A, Fig. 6, insets).

Lastly, the most valid comparison of 1-D and 2-D separations addresses not only resultant peak capacities, but also takes into account the overall execution time for the separations. A 2-D (SCX-RPLC) separation scheme was designed not to exceed the 200-min 1-D separation described earlier. Based on the data obtained during the 1-D analysis of the ribosomal protein mixture, the reversed-phase gradient was optimized by adjusting gradient range to 20–50% B over a period of 18 min (Fig. 5). Analysis of the SCX flowthrough fraction, and eight subsequent salt step fractions from the first dimension in the 2-D system yielded an apparent peak capacity of ~700, or ~3-fold higher than a 1-D separation conducted over the same overall run duration.

3.8. LC/MS detection of proteins from on-line separations

The inability to resolve components that differ little in 1-D HPLC retention will have a range of effects depending on detector specificity, detector resolution, and experimental objectives. For example, quantitative analysis of a protein mixture using UV, fluorescence, or RI detection requires effective component resolution for generation of meaningful data. In contrast, ESI mass spectrometry is a more forgiving mode of detection that provides an addi-

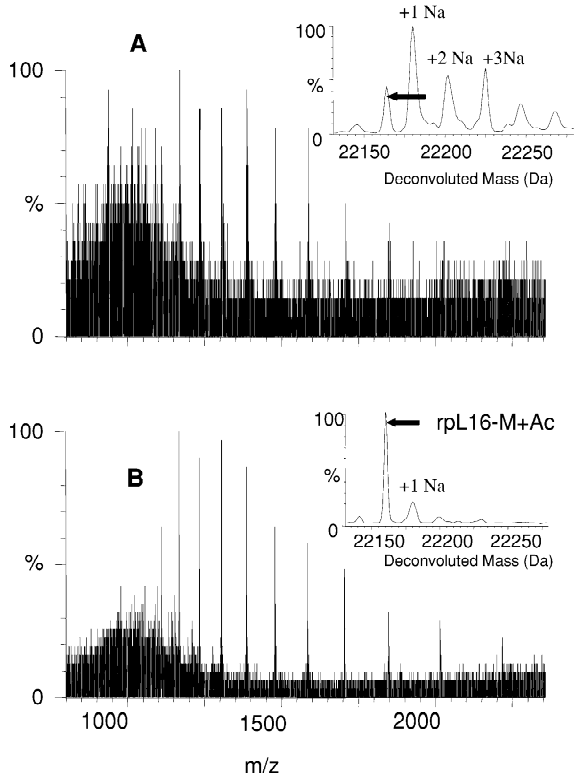

Fig. 6. Effect of diverting the non-volatile components from the first dimension SCX separation. An enriched ribosomal fraction was analyzed by SCX-RPLC–ESI–TOF MS using conditions identical to Fig. 5. Diversion of the first seven column volumes of the reversed-phase gradient to waste resulted in complex MS spectra (panel A), that upon deconvolution (panel A, inset) revealed a component eluting at 113.2 min into the run (black arrow) along with a series of adduct peaks. Increasing the diversion time to 14 column volumes revealed simplified mass spectra (panel B) and deconvoluted mass spectra (panel B, inset) for this component (elution time 113.4 min) now identifiable as yeast large ribosomal subunit L16B (acetylated and lacking the initiating methionine; expected mass 22 160.1, observed mass 22 160.5).

tional effective dimension of separation (i.e. ions in m/z space). Nevertheless, charge competition at the electrospray interface will create bias toward better ionizing species, and skew the detected population so that it is no longer representative of the analyte population present in the LC effluent. This is particularly true when coeluting components vary significantly in concentration (i.e. limitations on the dynamic range of the detector), or in chemical nature. In simplest terms, one could easily imagine mass spectral suppression of acidic and neutral proteins in the presence of equi-abundant or even lower abundance basic proteins. When processing protein ESI data, an additional concern is that the resultant spectral quality affects the ability to produce interpretable deconvoluted spectra for component identification.

The choice of ESI–TOF MS as a detector for the on-line 1-D and 2-D separations was made to provide the additional resolution necessary to resolve overlapping charge state envelopes of multiple coeluting components during the 1-D separation, while providing the mass accuracy (~20–50 ppm using external calibration) to make tentative identifications of eluted components. In practice, the use of quadrupole based mass detection systems would place greater demands on a separation, while high field ICR based detectors place less of a demand, relative to the high resolution TOF analyzer used in these studies.

LC/ESI–TOF MS experiments using lysozyme demonstrated that component identifications using maximum entropy deconvolution could accurately identify components barely obvious from total ion chromatograms (Fig. 7). As can be seen from this figure, 4 pmol of lysozyme (~440 fmol split to MS) produced interpretable deconvoluted spectra despite the small S/N (<2) of the peak on the TIC.

The more complicated yeast ribosomal sample showed that even with a total protein loading of only 750 ng, interpretable LC/MS mass spectra could be acquired (Fig. 8A). Maximum entropy deconvolution of a peak (RT: 8.6 min) from these 1-D analyses (Fig. 8B) corresponded to acetylated ribosomal subunit rpS11A lacking the initiating methionine (rpS11AB−M+Ac, 17 659.8 Da), which can be observed at the lowest injected amount (0.75 μg). A second component (rpL26A−M, 14 102.6 Da) is also observed which comprises the major component of the previous peak (7.75 min) on the TIC. Assuming 0.75–1.5 μg of total protein load as a reasonable lower limit for this analysis (based on analysis of other components as well, data not shown), and conservatively assuming 100 equimolar components with an average MW of 17 700, we believe we are reliably detecting ribosomal proteins at a load of 0.4–0.9 pmol per component, with roughly 50–100 fmol/component split to the mass spectrometer for

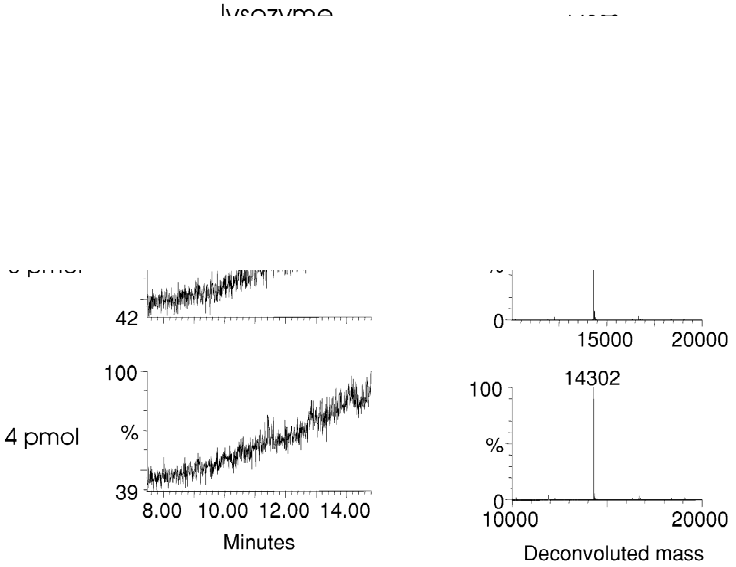

Fig. 7. MS sensitivity of the reversed-phase separation process using TFA as the acid modifier. An injection of 40, 8 and 4 pmol of lysozyme was made onto a 50×2.1 mm Symmetry300 C_4 column. The column was eluted with a 20–50% B gradient over 18 min at 0.5 ml/min, where eluent A=0.1% TFA in water and eluent B=0.1% TFA in acetonitrile. Detection was performed by positive ion ESI–MS-TOF scanning from 500 to 3000 *m/z*.

analysis. This apparent 4- to 10-fold increase in apparent increased MS sensitivity for ribosomal proteins versus lysozyme may be due to reduced sample losses when analyzing complex mixtures rather than single components, as well as issues relating to the relative behavior of individual proteins during separation and the electrospray ionization processes.

3.9. LC/ESI–TOF MS analysis of an enriched yeast ribosomal protein fraction

As described above, one argument at the heart of employing multidimensional chromatography versus a single mode of separation is not that the approach generates higher theoretical peak capacity, but rather that the additional mode of separation resolves components not otherwise resolved by either individual separation mode due to chemical and/or structural similarity. The spectra in Fig. 9 demonstrate the utility of coupling 2-D chromatography with ESI-MS detection for analysis of the enriched yeast ribosomal protein sample. During a high resolution 1-D HPLC separation of the ribosomal protein mixture (Fig. 4), the peak at 86.47 min (MS scans 6730–6777) contained at least three co-eluting components (masses of 12 821.80, 14 447.6, and 15 239.8) in the mass range of ribosomal proteins (3–46 kDa) as evidenced by the mass spectra (Fig. 9A) and deconvoluted mass spectra (Fig. 9B) summed over the full width of the peak. During the 2-D experiment, the two identifiable ribosomal subunits (rpS24AB−M+Ac, RT: 110.12 min, mass observed 15 240.5 Da, mass predicted 15 239.7; and rpL31A−M, RT: 68.71 min, mass observed 12 822.4, mass predicted 12 822.0), were resolved from each other into differing salt elution steps, as well as from a non-identified component (RT: 48.83, mass observed 14 448.5 Da) that elutes in the earliest step elution from SCX (Table 2). Fig. 9C,D (MS scans 8572–8580) and Fig. 9E,F (MS scans 5348–5362) from the 2-D run clearly demonstrate the simplification of the raw and deconvoluted spectra from the two previously co-eluting ribosomal subunits. As would be predicted from coelution during the 1-D separation, the relative reversed-phase elution times of rpS24AB, rpL31A, and the unidentified component in the reversed-phase dimension of the 2-D separation are very similar (8.74, 8.71, and 8.83 min, respectively).

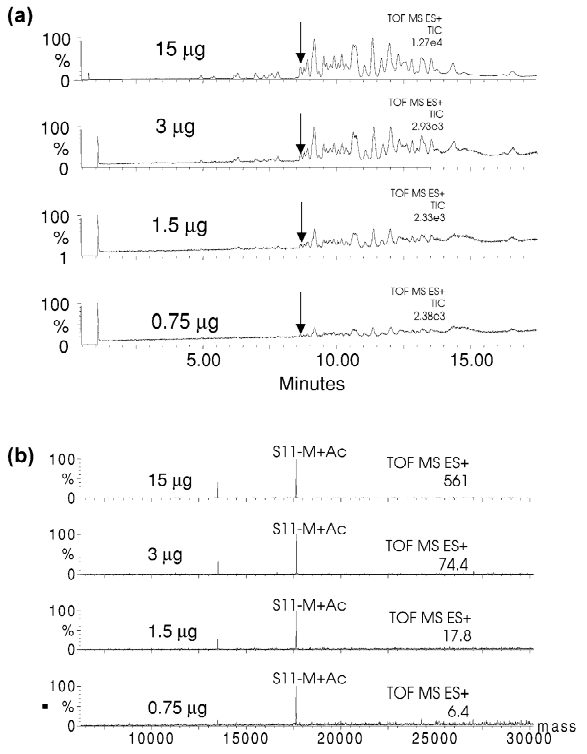

Fig. 8. The extracted deconvoluted spectra of a ribosomal sample injected in four different concentrations. Injections of 15, 3, 1.5 and 0.75 μg, respectively, were made onto a 50×2.1 mm Waters Symmetry300 C_4 column. The column was eluted with a 20–50% B gradient over 18 min at 0.5 ml/min, where eluent A=0.1% TFA in water and eluent B=0.1% TFA in acetonitrile. Detection was performed by positive ion ESI–MS-TOF scanning from 500 to 3000 *m*/*z*. The peak shown in the TIC traces (Fig. 8A) eluted from the reversed-phase system with a retention time of 8.5 min. Deconvoluted mass spectra for the peak are shown in Fig. 8B.

The resulting dataset from LC-LC/MS analysis of this complicated mixture is a list of observed masses that can be correlated to overall retention time of the experiment (or MS scan number), or relative retention time within either dimension. The separation scheme used in this work (step SCX elution, followed by short RPLC gradients) can be most clearly displayed by presenting an ordered set of observed masses organized first by SCX elution step and then by relative retention time in the second dimension (Table 2).

To generate the data in Table 2, the raw mass 2-D spectra from the 200-min SCX-RPLC–ESI–TOF MS analysis were deconvoluted using a maximum entropy deconvolution algorithm as described in the Experimental section. This procedure identified ~160 individual components that included several components apparently eluting in multiple ion-exchange fractions. Accounting for these cases, 157 (126 unique) masses ranging from 6629 to 43 731 Da were obtained from the analysis of an enriched yeast ribosomal fraction.

As can be seen in Tables 2 and 3, tentative assignments of 91 unique masses have been made to 66 known ribosomal proteins (80 isoforms), with an additional 35 of the 126 unique masses not yet assigned. Assignments were produced by searching a list of ribosomal subunit average masses calculated from translated yeast open reading frame sequences, assuming either loss or retention of the initiating methionine and presence or absence of N-terminal acetylation. Several additional assignments were made by applying known methylations, phosphorylations and proteolytic maturations described in the yeast protein database [25,26], and additional primary sources [22,37–48]. For the 2-D ribosomal dataset (Tables 2 and 3), assignment of ~80% of ribosomal subunits (70% of all potential ribosomal subunit isoforms) was accomplished with an average absolute mass error of ~50 ppm (SD~25 ppm). Table 3 also demonstrates that this approach produced essentially equal coverage of proteins from both the large and small subunits of the yeast ribosome. Achieving mass measurement accuracy (MMA) in the range of 20–50 ppm using same-day external calibrations is routinely accomplished by our laboratory, and efforts to employ internal on-line lock mass calibrations should produce further MMA improvements to aid in distinguishing subunits with closely related masses.

It can be reasonably argued that measurements of intact protein mass (even at low ppm MMA) by themselves are insufficient to conclusively identify proteins, particularly given that most proteins are modified in some manner from the predicted sequence. This challenge effectively grows at least linearly with the number of proteins in a database, and combinatorially with the number of possible modifications. Of course, this argument assumes no pre-knowledge of a sample, which is often not the case when analyzing established biochemical path-

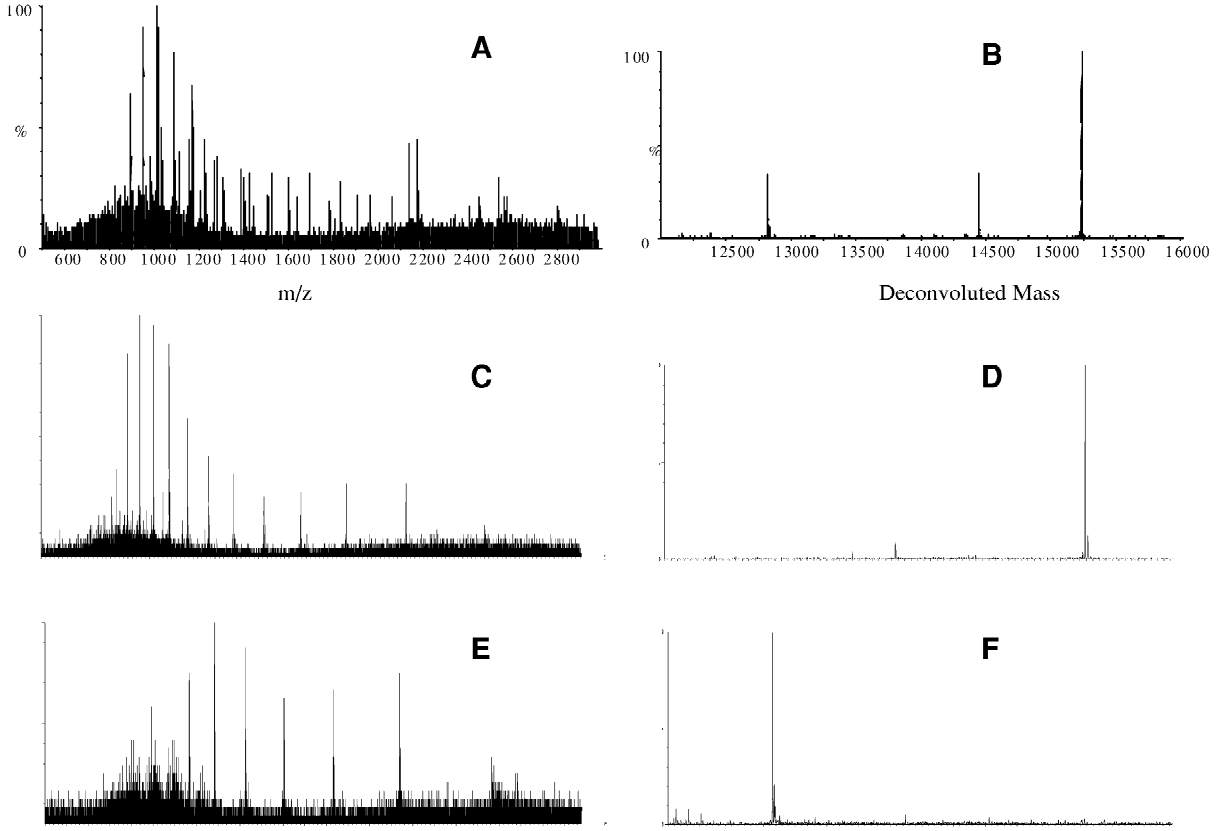

Fig. 9. The application of multidimensional chromatography to an enriched yeast ribosomal protein fraction results in increased mass spectral quality over a single dimension separation of the same overall duration. Yeast ribosomal proteins (70 μg) analyzed by 1-D RPLC-MS produced a peak of total ion current 86.47 min into the LC run (MS scans 6730–6777) that corresponded to a complex mass spectrum (panel A) that upon maximum entropy deconvolution (panel B) revealed the co-elution of at least three components within the mass range of ribosomal proteins (3–46 kDa). Application of a 2-D SCX-RPLC-MS separation to this sample resolves all three components. The ribosomal subunit rpS24AB elutes in the fifth SCX salt step (overall RT: 110.12 min) and produces spectra (panels C and D) consistent with a single component consistent with the acetylated form of the protein lacking the initiating methionine (Table 2), while rpL31A elutes in the third SCX step (overall RT: 68.71 min) with mass spectra (panels E and F) consistent with loss of the initiating methionine (Table 2).

ways and protein complexes. Furthermore, databases such as SGD [49] and YPD [25,26] now comb the scientific literature for evidence of potential modifications (from direct experimentation or by protein/domain/motif homology) and biomolecular interaction (e.g. co-IPP, 2-hybrid, co-regulation) that can aid in interpretation of complicated intact protein mixture data.

In reality, the need to confirm identifications using a secondary method will often be required for definitive identification of components. Several approaches, including the use of stable isotope labeling [50,51] and gas phase dissociation [52,53] of intact proteins have been proven in principle to permit protein identification from intact mass data. Using the system described here we are effectively diverting ~88% of the second dimension effluent for UV detection and fraction collection, where more commonplace approaches to protein identification from proteolytic digests (i.e. peptide mass fingerprinting and MS/MS sequencing [17,54–58]) can be applied. It is this interplay where definitive identifications from digested peptides are combined with accurate intact mass analysis (with modification information) that effectively permits relevant biological questions to be asked of complicated bio-

Table 2
SCX-RP–ESI-TOF-MS analysis of an enriched ribosomal protein fraction

IEX[a]	RP[b]	Observed[c]	Expected[d]	Delta[e]	ID[f]	Notes[g]
1	6.99	9788.5	9787.9	61	rpS21A + Ac	
1	7.39	9802.5	9802.0	51	rpS21B + Ac	
1	8.13	7607.0	7606.9	13	rpS28B + Ac	
1	8.13	7634.0	7633.9	13	rpS28A + Ac	
1	11.27	12 364.5			U	
1	11.27	12 420.5			U	
1	11.70	11 131.0	11 130.2	40	rpP2B + (P)	
1	11.70	14 686.5			U	
1	11.70	23 792.5			U	
1	12.06	6730.5			U	
1	12.06	10 826.0	10 825.9	14	rpP2A + (P)	
1	13.95	27 875.0	27 873.5	54	rpS0B − M + Ac	
1	13.95	27 938.0	27 935.5	89	rpS0A − M + Ac	
1	16.22	17 382.0			U	
1	16.88	10 658.5			U	
1	18.37	15 383.5	15 382.6	59	rpS12 − M + Ac	
2	3.34	6597.0	6596.6	61	rpS29B − M	
2	5.24	14 104.5			U	
2	7.19	8762.5	8762.3	25	rpS27A − M + Me	
2	8.83	11 046.5			U	
2	8.83	14 448.5			U	
2	8.83	15 761.0			U	
2	8.83	15 786.0	15 786.0	0	rpS19A − M	DA3
2	8.83	19 632.0	19 630.6	70	rpL11A − M + Ac	
2	8.83	19 661.0	19 660.7	18	rpL11B − M + Ac	
2	9.06	14 440.5	14 440.1	28	rpL23AB − M + 7 Me	
2	9.39	13 819.0	13 818.0	72	rpS8 − M + Ac or rpL35AB − M + Ac	
2	10.07	15 460.0			U	FS3
2	10.07	15 759.5	15 759.0	32	rpS19B − M or rpS16B − M + Ac or rpL25	NIB(3,5,6,7,8)
2	11.00	17 776.5	17 775.8	42	rpL12AB − M + 6 Me	
2	11.25	13 563.0	13 562.3	52	rpL22A − M	
2	11.25	21 546.0	21 545.2	37	rpS7B − M + Ac	
2	11.52	14 495.5	14 495.0	34	rpS22AB − M	FS3
2	11.83	14 695.0			U	
2	11.83	20 397.5			U	
2	11.83	21 534.5	21 533.2	60	rpS7A − M + Ac	
2	13.42	26 386.5	26 385.8	25	rpS3 + Me	
2	15.66	24 951.0	24 949.6	56	rpS5 − M + Ac	
2	15.66	30 443.0			U	
3	4.17	6529.5	6529.0	0	rpS29A − M	
3	5.04	14 103.5	14 102.6	64	rpL26A − M	
3	7.57	8695.5	8695.4	12	rpL38 − M	DP (DA/DS3)
3	7.57	12 023.5	12 023.0	42	rpL33A − M	
3	7.80	8695.5	8695.4	12	rpL38 − M	DP (DA/DS3)
3	7.80	12 037.0	12 037.0	0	rpL33B − M or rpS25A	
3	8.53	20 306.5			U	
3	8.71	10 897.5	10 898.8	123	rpP1A − M + Ac + (P)	
3	8.71	12 822.5	12 822.0	39	rpL31A − M	
3	8.71	12 836.5	12 836.0	39	rpL31B − M	
3	9.25	15 914.0	15 912.8	75	rpS15 − M + Ac	
3	9.25	19 856.5	19 855.4	55	rpL6B − M	
3	9.25	19 874.5	19 872.5	101	rpL6A − M + Ac	
3	9.58	15 786.0	15 786.0	0	rpS19A − M	DA2
3	9.83	15 064.5	15 064.0	33	rpL14B − M + Ac	

Table 2. Continued

IEX[a]	RP[b]	Observed[c]	Expected[d]	Delta[e]	ID[f]	Notes[g]
3	9.83	15 460.0			U	FS2, CO5
3	9.83	15 759.5	15 759.0	32	rpS19B−M or rpS16B−M+Ac or rpL25	NIB(2,5,6,7,8)
3	10.08	15 079.0	15 078.0	66	rpL14A−M+Ac	
3	11.06	22 169.0	22 167.6	63	rpS9B−M	FS4
3	11.38	14 495.0	14 495.0	34	rpS22AB−M	FS2
3	11.75	21 569.5	21 569.2	14	rpL9A or rpL9B−M+Ac	
3	12.19	18 898.5	18 897.5	53	rpL6B−M+Ac	
3	13.21	27 216.0			U	
3	13.21	27 362.5	27 360.8	62	rpS2−M+Ac	
3	14.57	33 586.5	33 583.9	77	rpL5−M or P0−M	NIB(4,5,6,7,8)
4	7.37	18 112.0	18 111.1	50	rpL21A−M	
4	7.37	18 143.5	18 143.2	17	rpL21B−M	
4	8.98	20 422.0	20 420.4	98	rpL17B−M or rpL17A−M	DP (DA/DS4)
4	9.10	20 420.0	20 420.4	5	rpL17B−M or rpL17A−M	DP (DA/DS4)
4	9.47	11 937.0	11 936.2	71	rpS25A−M+2 Me	
4	9.47	19 832.0	19 830.5	76	rpL6A−M	
4	9.81	15 627.5	15 626.5	64	rpL25−M	
4	10.03	29 281.0	29 279.2	61	rpS4AB−M	FS5
4	11.39	22 168.5	22 167.6	41	rpS9B−M	FS3
4	12.17	16 949.0	16 948.5	30	rpS18AB−M+Ac	CO8
4	12.36	16 899.0	16 897.8	71	S13−M	
4	12.81	17 525.0	17 524.5	29	rpL24A−M+Ac	
4	12.81	27 984.0	27 980.8	114	rpL8B−M	
4	12.81	27 999.0			U	
4	14.06	27 509.0	27 507.2	65	rpL7A−M	CO(6,8), FS(5)
4	14.80	33 587.0	33 583.9	92	rpL5−M or rpP0−M	NIB(3,5,6,7,8)
5	6.64	17 660.5	17 659.8	40	rpS11AB−M+Ac	
5	6.97	15 401.0	15 400.2	52	rpL27A−M	
5	7.43	11 317.5			U	
5	7.43	15 507.0			U	
5	7.43	15 961.0	15 959.4	100	rpS19A+Ac	
5	7.43	27 279.0	27 277.4	59	rpL2AB−M	
5	8.01	11 004.5	11 004.1	36	rpL36B−M	
5	8.61	13 779.5	13 778.5	73	rpL35AB−M	CO7
5	8.74	15 240.5	15 239.7	52	rpS24AB−M+Ac	
5	8.97	9960.0	9959.8	20	rpL43AB−M	
5	8.97	15 658.0	15 657.3	45	rpS17A−M	
5	8.97	15 674.5	15 672.2	115	rpS17B−M	
5	9.95	15 459.0			U	CO3
5	9.95	15 758.5	15 759.0	32	rpS19B−M or rpS16B−M+Ac or rpL25	NIB(2,3,6,7,8)
5	9.95	29 280.5	29 279.2	44	rpS4AB−M	FS4
5	10.55	26 927.0			U	
5	10.55	26 998.5	26 996.5	74	rpS6AB	
5	10.77	21 758.5			U	
5	10.77	22 113.5	22 112.1	63	rpL16A−M+Ac	
5	11.35	22 162.0	22 160.1	86	rpL16B−M+Ac	
5	11.94	38 976.0	38 973.2	72	rpL4B−M+Ac	
5	11.94	39 006.0	39 003.2	72	rpL4A−M+Ac	
5	12.30	25 232.0	25 230.4	63	rpL10−M	CO7
5	13.30	28 613.0	28 612.4	21	rpS1A−M	
5	13.30	28 681.0	28 681.5	17	rpS1B−M	
5	14.05	27 509.0	27 507.2	65	rpL7A−M	CO(7), FS(4,6)
5	14.70	33 585.0	33 583.9	33	rpL5−M or rpP0−M	NIB(3,4,6,7,8)
6	7.07	13 510.0	13 510.0	37	rpL34B−M	DP (DA6)
6	7.31	13 509.0	13 510.0	74	rpL34B−M	DP (DA6)
6	7.98	12 278.5			U	
6	7.98	22 395.0	22 394.0	45	rpL13B−M	

Table 2. Continued

IEX[a]	RP[b]	Observed[c]	Expected[d]	Delta[e]	ID[f]	Notes[g]
6	8.65	14 641.0	14 640.2	55	rpL32−M	
6	9.38	16 276.5			U	
6	9.38	16 591.5	16 590.5	60	rpL28	FS7
6	10.13	29 319.5	29 321.2	58	rpS4AB−M+Ac	
6	10.14	15 759.5	15 759.0	32	rpS19B−M or rpS16B−M+Ac or rpL25	NIB(2,3,5,7,8)
6	10.14	29 295.0			U	
6	10.14	31 518.5			U	
6	10.53	43 668.0	43 668.8	18	rpL3−M+Ac	FS7
6	10.53	43 709.0			U	
6	14.14	27 507.0	27 507.2	7	rpL7A−M	CO(4,8), FS(5,7)
6	14.14	27 527.5			U	CO8
6	14.91	33 586.0	33 583.9	63	rpL5−M or rpP0−M	NIB(3,4,5,7,8)
7	3.16	12 108.5	12 108.5	4	rpL42AB−M+ 2 Me	
7	4.38	17 548.0	17 547.6	23	rpL24B	
7	4.53	17 614.5	17 613.7	45	rpL24A	
7	6.14	15 937.5			U	DP (DA/DS7)
7	6.69	15 939.5			U	DP (DA/DS7)
7	8.02	13 374.0	13 373.7	22	rpS26A−M	
7	8.42	22 359.5	22 358.6	40	rpS8AB−M	
7	8.69	13 779.0	13 778.5	36	rpL35AB−M	CO5
7	9.41	16 591.0	16 590.5	30	rpL28	FS6
7	9.41	16 607.0			U	
7	9.97	15 759.0	15 759.0	0	rpS19B−M or rpS16B−M+Ac or rpL25	NIB(2,3,5,6,8)
7	10.31	43 668.0	43 668.8	18	rpL3−M+Ac	FS6
7	11.14	20 433.0	20 432.1	44	rpL18AB−M	
7	11.14	20 450.5			U	CO9 FS8
7	12.02	19 509.5			U	
7	12.38	16 914.5			U	
7	12.38	25 229.0	25 230.4	55	rpL10−M	CO5
7	13.34	28 699.5			U	
7	14.06	27 507.0	27 507.2	7	rpL7A−M	CO(5), FS(6,8)
7	14.74	33 585.0	33 583.9	33	rpL5−M or rpP0−M	NIB(3,4,5,6,8)
8	3.75	8677.0	8677.4	50	Processed rpS31	FS9
8	7.79	21 574.5	21 573.1	65	rpL19AB−M	FS9
8	10.19	15 758.5	15 759.0	32	rpS19B−M or rpS16B−M+Ac or rpL25	NIB(2,3,5,6,7)
8	11.28	20 449.5			U	FS(7,9)
8	12.20	16 948.0	16 948.5	30	rpS18AB−M+Ac	CO4
8	14.10	27 508.0	27 507.2	29	rpL7A−M	CO(4,6), FS(7)
8	14.10	27 527.5			U	CO6, FS9
8	14.87	33 586.5	33 583.9	77	rpL5−M or P0−M	NIB(3,4,5,6,7)
9	3.79	8677.0	8677.4	50	Processed rpS31	FS8
9	7.50	21 575.0	21 573.1	88	rpL19AB−M	FS8
9	7.84	24 293.0	24 291.2	74	rpL15AB−M	
9	8.24	24 334.0	24 333.2	33	rpL15AB−M+Ac	
9	11.17	20 450.0			U	CO7 FS8
9	14.07	27 525.0			U	FS9

[a] SCX step fraction: due to experimental construction, fraction 1 contains material not retained on the SCX column that bound the RP column under initial binding conditions, while fraction 0 (0–20 min of gradient) served as a blank run, while the first RP column was loaded.

[b] Relative reversed-phase LC gradient time measured from the start of each RP gradient, rather than the start of the overall 2-D experiment.

[c] Neutral mass determined by application of a maximum entropy algorithm using a 0.5-Da output bin size.

[d] Average neutral mass of ribosomal subunit(s) calculated from yeast genomic ORF data.

[e] Absolute difference in PPM between observed and expected ribosomal subunit masses.

[f] Ribosomal subunit identification using the standardized nomenclature [32,33], except that isoforms (A and B) of a subunit with complete identity are grouped as ''AB''. ''U'' listed for unidentified components. Modifications: Ac, acetylation; −M, loss of methionine; Me, methylation; (P), phosphorylation; Processed, mature proteolytic fragment.

[g] Notes relating to chromatographic behavior (COx, RP carry-over with fraction x; DP, doublet peak of component that is seen twice in same fraction; FSx, IEX fraction splitting of component with fraction x) or possible additional modifications (DA/DSx, possible deamidation or disulfide related species of component in fraction x; DAx, possible deamidation related species of component in fraction x). NIB, or near isobaric is listed when an observed mass corresponds to more than one ribosomal subunit within experimental error.

Table 3
Coverage chart of ribosomal components observed during SCX-RP–ESI–TOF MS analysis of an enriched yeast ribosomal fraction

	Small	Large	Acidic	Total
Possible[a]				
Total ORFs	56	76	5	137
Unique ORFs	47	64	5	116
Subunits	32	43	3	78
Observed[b]				
Unique ORFs	33	43	4	80
Subunits	27	34	3	66
% Unique ORFs observed	70	67	80	70
% Subunits observed	84	79	100	82

[a] Total number of canonical [32,33] ORFs, unique ORFs (by sequence), and ribosomal subunits in the yeast small, large, and acidic structural subunits.

[b] Number of unique ORFs, and ribosomal subunits in each structural subunit observed during the 2-D SCX-RP–ESI–TOF MS analysis of the enriched ribosomal fraction presented in Fig. 5.

logical samples. The importance of effective separations to generate fractions for digestion and peptide mass fingerprinting, infusion MS/MS, or LC/MS/MS sequencing should not be understated, as the complexity presented to an MS detector will ultimately dictate separation requirements and post-run processing efforts required for full characterization of the components.

An illustration of the power of coordinating intact and peptide digest information is provided by MALDI–TOF MS analysis of one of the collected fractions. As described previously, only 10% of RP column effluent is directed to the ESI-MS for intact protein analysis, while the vast majority is diverted for fraction collection. Fig. 10 shows a detailed analysis of fraction 35, which was collected during a 2-min period during the development of the RP gradient from the third IEX step fraction from the experiment shown in Fig. 5. During this elution time window, four TIC peaks were observed (Fig. 10A). The summed spectra for this entire window yielded a complex MaxEnt deconvoluted spectrum (Fig. 10B) that contained masses corresponding to several large ribosomal subunit proteins (rpL38, rpL31A, rpL31B, rpL33A, and rpL33B). In addition, a component with an unassigned mass (20 307.0 Da) was observed that did not correspond to a known ribosomal protein with or without N-terminal processing, or any modifications described in the literature.

The fraction collected during this elution window was dried, reduced, digested with trypsin, and spotted for MALDI peptide mass fingerprinting as described in the Experimental section. The PMF analysis (Fig. 10C) yielded confirmation of both isoforms of rpL31 and rpL33, but not of the smaller rpL38 protein. This observation is not unexpected, as the smaller number of potentially identifiable fragments arising from small highly basic proteins like rpL38 (8.8 kDa, pI~11.07) would favor detection by intact protein analysis methods. The confirmation of the intact mass data for rpL31 and rpL33 isoforms by PMF also confirms the absence of additional modifications, the processing of the N-terminal methionine, and the lack of acetylation on the penultimate alanine residues of both subunits. The PMF data also revealed high sequence coverage (59%) of rpL20A/B, for which no intact mass data had been obtained. This prompted us to further examine the unassigned intact mass, and sequence of rpL20 to conclude the unidentified mass corresponds to a truncated form of rpL20 lacking three (MYL, rpL20B) or nine (MKILVILSV, rpL20A) amino acids from the N-terminus. Surprisingly, these regions contain all differences between the two isoforms, and the ''processed'' fragment observed by intact mass is identical with both subunit isoforms. The biology behind this instance of isoform differentiation, and potential processing mechanisms that led to observation of a common processed structure remain to be elucidated.

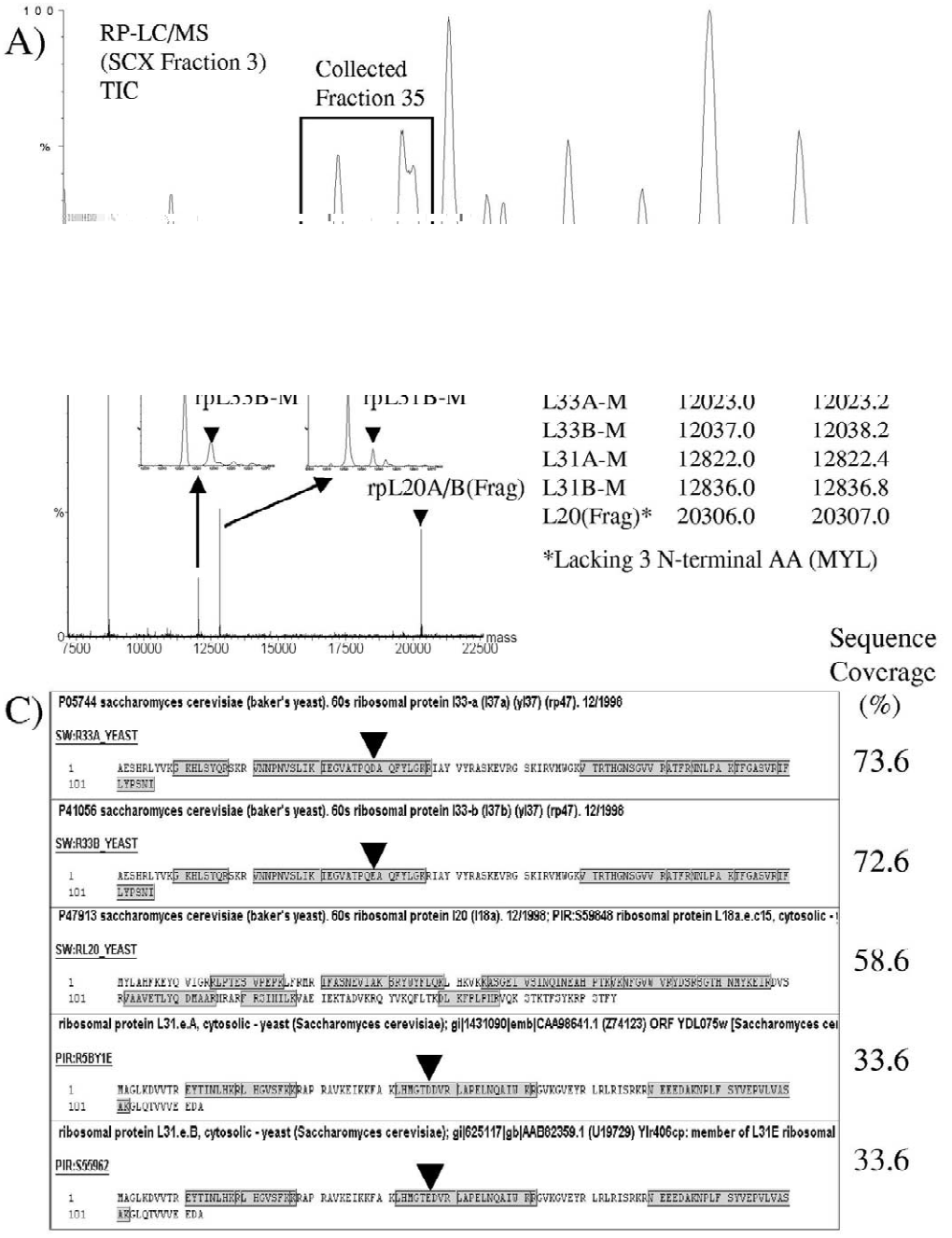

Fig. 10. Detailed analysis of collected fraction 35. The expanded view for the reversed-phase gradient LC/MS of the third IEX fraction (A) shows a region collected as fraction 35 encompassing four TIC peaks (shown in the box). The effluent was split with 10% analyzed by intact protein ESI–TOF MS and the remaining 90% collected for tryptic digestion and analysis peptide mass fingerprinting. The deconvoluted mass spectra of this region (B) reveals at least six components, five of which can be assigned to known ribosomal proteins (rpL38, rpL31A/B, and rpL33A/B) modified only by removal of the initiating methionine, and one unassigned mass (20 307.0 Da). Peptide mass fingerprinting data for this fraction (C) confirms assignments for the assigned subunits, and led to the identification of a truncated rpL20A/B, with the intact mass consistent with subunits lacking several residues from the N-termini. Arrows point to peptides that permit protein isoform characterization.

3.10. Quality of 2-D LC–LC/MS separations

The data from the LC/MS analysis are also an important element in assessing the chromatographic performance of a 2-D separation. Both TIC and UV response indicate at least 107 peaks are resolved or partially resolved by the system, far in excess of the 53 peaks resolved or partially resolved in the 1-D separation. However, neither UV nor TIC response alone or in tandem is capable of identifying peaks appearing in more than one ion-exchange fraction that are due to a single protein structure. Only from deconvoluted MS spectra can individual masses that appear in multiple fractions be recognized. In those cases where this occurs, the assignment of such signals to the same protein is confirmed by consistent retention times in the reversed-phase separation.

A major concern with protein separations, particularly with multi-dimensional methods, is the possibility that a single molecular species is observed in multiple fractions. This phenomenon complicates data analysis and can lead to erroneous data interpretation as to the number of distinct entities separated. There are potentially two major causes for multiple occurrences of the protein reidentification. First, if a protein elutes at the end of an ion-exchange step, some of mass may remain bound on the column to be eluted in the subsequent step. We have attempted to minimize this ''fraction splitting'' by implementing a step-gradient protocol rather than the more usual continuous gradients used for protein chromatography. This fraction splitting phenomenon is characterized by the appearance of the same mass at comparable reversed-phase retention time in consecutive ion-exchange fractions. Of the 91 tentatively identified ribosomal components, 67 were isolated in a single ion-exchange fraction. This minimal overlap between fractions is clear vindication for the step gradient approach to coupling ion-exchange chromatography with a reversed-phase step. However, even with the step gradient operation in the ion-exchange mode, several incidences of fraction splitting were observed. Examples (Table 2) include rpS31 in fractions 8 and 9, rpL28 in fractions 6 and 7, S9B−M in fractions 3 and 4 and rpS4AB−M in fractions 4 and 5. In all four examples much weaker signals are observed in the later eluting fraction, indicating that the bulk of the protein mass eluted in the earlier fraction.

Another potential mechanism for the appearance of components in multiple fractions is carry-over from the reversed-phase step. This well-described phenomenon [59] has been attributed to protein remaining on the stationary phase following a gradient run and appearing in subsequent analyses even if no further injection is made. As our system distributes the ion-exchange step elution to alternating reversed-phase columns, carry-over will be characterized by the appearance of a particular mass in fraction n followed by its appearance in fraction $n+2$. Proteins exhibiting carry-over behavior (Table 2) include rpL35AB−M (nominal mass 13 779) and rpL10−M (nominal mass 25 229), both observed in fractions 5 and 7.

In several cases the same nominal mass is observed in consecutive fractions, but with significantly different ($\Delta t > 0.20$ min) reversed-phase retention times. These are not likely to be identical protein structures, although they may be closely related structures such as those derived from deamidation or other minor modifications including disulfide bond reduction that cause only a small change in mass. The mass accuracy of the LC–TOF analysis may be insufficient to distinguish such minor changes in structure. Again, detailed analysis of peptides derived from digests of the collected fractions will reveal the nature of the relationship between these components.

Only six masses (nominally 15 460, 15 759, 20 450, 27 509, 27 525 and 33 585) out of a total greater than 120 (<5% of the total) appear in more than two fractions. Three of these, with nominal masses of 15 460, 20 450 and 27 525, are unassigned (labeled ''U'' in Table 2). Mass 15 759, which is consistent with at least three nearly isobaric subunits (rpS19B−M, rpS16B−M+Ac and rpL25), appears over a wide range of reversed-phase retention times (9.83–10.19 min). Similarly, nominal mass 33 585 (assigned as rpL5−M or rpP0−M) appears in six fractions, but may arise from two nearly isobaric proteins. In addition, variations in reversed-phase retention times are consistent with several distinct structures. Only one assigned ribosomal protein, rpL7A−M with nominal mass 27 509, appears in more than two IEX fractions with essentially the same retention in the reversed-phase step for each of those fractions.

4. Conclusions

We have presented a practical multi-dimensional chromatographic system with on-line MS detection and off-line fraction collection for the analysis of protein mixtures. Detection with an ESI–TOF instrument provided good MMA in the range of 20–50 ppm. A detailed analysis of a yeast ribosomal protein preparation with the 2-D LC–LC/MS system allowed tentative identification of >80% of the subunits in the complex. The configuration with two reversed-phase columns for the second dimension separation provided two major benefits over a single column configuration. First, the overall analysis time with two columns was reduced by at least 30% due to the efficiency of off-line loading and re-equilibration steps. Also significant was the ability to discriminate between elution problems occurring in the IEX step (fraction splitting) and those occurring in the second step (carry-over).

The value derived from applying two orthogonal separation modes increases proportionally to sample complexity, but the benefits also extend to samples such as the ribosomal protein fraction with intermediate complexity but highly similar proteins (size, p*I*, common cellular molecular environments, etc.). As currently implemented, we believe protein samples of much greater complexity and diversity can be successfully analyzed using the 2-D LC/ESI-MS approach presented here. Current experimental results do not, however, completely describe the limitations of the approach. Ribosomal proteins are, on average, more hydrophilic than many proteins, especially membrane-associated species, and more work is required to delineate the applicability to hydrophobic proteins where issues of recovery from either the IEX or RP columns could be problematic. In addition, the ribosomal proteins only span a limited mass range, the largest identified structure being only 43.6 kDa. We can anticipate a much higher useful range, though, as Lubman and coworkers, for example, have reported on-line ESI–TOF MS results for proteins up to 85 kDa [8].

Lastly, although we can only tentatively identify proteins and potential modifications using the intact mass data, it is in conjunction with more conventional peptide analysis techniques including peptide mass fingerprinting or CID MS/MS sequence analysis that this system will provide maximum information to a researcher. This is in contrast to global digestion procedures [1,2] where loss of information regarding the intact proteins makes identifying post-translational modifications or co-translational protein processing difficult. The system as currently configured provides a potentially viable alternative to 2-D gels that can be implemented using commercially available technology to acquire intact protein mass data, identify the component(s) with peptide based identification strategies, and provide a rational basis to examine the existence of specific protein modifications.

References

[1] A.J. Link, J. Eng, D.M. Schieltz, E. Carmack, G.J. Mize, D.R. Morris, B.M. Garvik, J.R. Yates III, Nat. Biotechnol. 17 (1999) 676.
[2] S.P. Gygi, B. Rist, S.A. Gerber, F. Turecek, M.H. Gelb, R. Aebersold, Nat. Biotechnol. 17 (1999) 994.
[3] H. Gao, Y. Shen, T.D. Veenstra, R. Harkewicz, G.A. Anderson, J.E. Bruce, L. Pasa-Tolic, R.D. Smith, J. Microcol. Sep. 12 (2000) 383.
[4] M.T. Davis, J. Beierle, E.T. Bures, M.D. McGinley, J. Mort, J.H. Robinson, C.S. Spahr, W. Yu, R. Luethy, S.D. Patterson, J. Chromatogr. B Biomed. Sci. Appl. 752 (2001) 281.
[5] A. Butt, M.D. Davison, G.J. Smith, J.A. Young, S.J. Gaskell, S.G. Oliver, R.J. Beynon, Proteomics 1 (2001) 42.
[6] P.K. Jensen, L. Pasa-Tolic, K.K. Peden, S. Martinovic, M.S. Lipton, G.A. Anderson, N. Tolic, K.K. Wong, R.D. Smith, Electrophoresis 21 (2000) 1372.
[7] P.K. Jensen, L. Pasa-Tolic, G.A. Anderson, J.A. Horner, M.S. Lipton, J.E. Bruce, R.D. Smith, Anal. Chem. 71 (1999) 2076.
[8] D.B. Wall, M.T. Kachman, S.S. Gong, S.J. Parus, M.W. Long, D.M. Lubman, Rapid Commun. Mass Spectrom. 15 (2001) 1649.
[9] D.B. Wall, S.J. Parus, D.M. Lubman, J. Chromatogr. B Biomed. Sci. Appl. 763 (2001) 139.
[10] G.J. Opiteck, S.M. Ramirez, J.W. Jorgenson, M.A. Moseley III, Anal. Biochem. 258 (1998) 349.
[11] J.C. Giddings, J. High Resolut. Chromatogr. Chromatogr. Commun. 10 (1987) 319.
[12] J.C. Giddings, Anal. Chem. 56 (1984) 1258A.
[13] R.E. Murphy, M.R. Schure, J.P. Foley, Anal. Chem. 70 (1998) 1585.
[14] J.M. Davis, Anal. Chem. 63 (1991) 2141.
[15] M.M. Bushey, J.W. Jorgenson, Anal. Chem. 62 (1990) 161.
[16] G.J. Opiteck, J.W. Jorgenson, Anal. Chem. 69 (1997) 2283.
[17] B.E. Chong, F. Yan, D.M. Lubman, F.R. Miller, Rapid Commun. Mass Spectrom. 15 (2001) 291.

[18] K.K. Unger, K. Racaityte, K. Wagner, T. Miliotis, L.E. Edholm, R. Bischoff, G. Marko-Varga, J. High Resolut. Chromatogr. 23 (2000) 259.
[19] K. Wagner, T. Miliotis, G. Marko-Varga, R. Bischoff, K.K. Unger, Anal. Chem. 74 (2002) 809.
[20] K. Wagner, K. Racaityte, K.K. Unger, T. Miliotis, L.E. Edholm, R. Bischoff, G. Marko-Varga, J. Chromatogr. A 893 (2000) 293.
[21] M. Capel, D. Datta, C.R. Nierras, G.R. Craven, Anal. Biochem. 158 (1986) 179.
[22] S.-W. Lee, S.J. Berger, S. Martinovic', L. Pasa-Tolic', G.A. Anderson, Y. Shen, R. Zhao, R.D. Smith, Proc. Natl. Acad. Sci. USA 99 (2002) 5942.
[23] S.J.S. Hardy, C.G. Kurland, P. Voynow, G. Mora, Biochemistry 8 (1969) 2897.
[24] A. Goffeau, B.G. Barrell, H. Bussey, R.W. Davis, B. Dujon, H. Feldmann, F. Galibert, J.D. Hoheisel, C. Jacq, M. Johnston, E.J. Louis, H.W. Mewes, Y. Murakami, P. Philippsen, H. Tettelin, S.G. Oliver, Science 274 (1996) 546.
[25] M.C. Costanzo, J.D. Hogan, M.E. Cusick, B.P. Davis, A.M. Fancher, P.E. Hodges, P. Kondu, C. Lengieza, J.E. Lew-Smith, C. Lingner, K.J. Roberg-Perez, M. Tillberg, J.E. Brooks, J.I. Garrels, Nucleic Acids Res. 28 (2000) 73.
[26] P.E. Hodges, A.H. McKee, B.P. Davis, W.E. Payne, J.I. Garrels, Nucleic Acids Res. 27 (1999) 69.
[27] I. Mazsaroff, F.E. Regnier, J. Chromatogr. 443 (1988) 119.
[28] W. Kopaciewicz, M.A. Rounds, J. Fasnauh, F.E. Regnier, J. Chromatogr. 266 (1983) 3.
[29] C.G. Huber, A. Premstaller, J. Chromatogr. A 849 (1999) 161.
[30] A. Apffel, S. Fischer, G. Goldberg, P.C. Goodley, F.E. Kuhlmann, J. Chromatogr. A 712 (1995) 177.
[31] L.R. Snyder, in: C. Horvath (Ed.), High-Performance Liquid Chromatography, Advances and Perspectives, Academic Press, New York, 1980, p. 207.
[32] R.J. Planta, W.H. Mager, Yeast 14 (1998) 471.
[33] W.H. Mager, R.J. Planta, J.-P. Ballesta, J.C. Lee, K. Mizuta, K. Suzuki, J.R. Warner, J. Woolford, Nucleic Acids Res. 25 (1997) 4872.
[34] Y. Shen, N. Tolic, R. Zhao, L. Pasa-Tolic, L. Li, S.J. Berger, R. Harkewicz, G.A. Anderson, M.E. Belov, R.D. Smith, Anal. Chem. 73 (2001) 3011.
[35] Y. Shen, R. Zhao, M.E. Belov, T.P. Conrads, G.A. Anderson, K. Tang, L. Pasa-Tolic, T.D. Veenstra, M.S. Lipton, H.R. Udseth, R.D. Smith, Anal. Chem. 73 (2001) 1766.
[36] L. Tolley, J.W. Jorgenson, M.A. Moseley, Anal. Chem. 73 (2001) 2985.
[37] D. Becker-Ursic, J. Davies, Biochemistry 15 (1976) 2289.
[38] R.J. Arnold, B. Polevoda, J.P. Reilly, F. Sherman, J. Biol. Chem. 274 (1999) 37035.
[39] B.-U. Doris, J. Davies, Biochemistry 15 (1976) 2289.
[40] T. Kruiswijk, A. Kunst, R.J. Planta, W.H. Mager, Biochem. J. 175 (1978) 221.
[41] J. Lhoest, Y. Lobet, E. Costers, C. Colson, Eur. J. Biochem. 141 (1984) 585.
[42] Y. Lobet, J. Lhoest, C. Colson, Biochim. Biophys. Acta 997 (1989) 224.
[43] T. Naranda, J.P.G. Ballesta, Proc. Natl. Acad. Sci. USA 88 (1991) 10563.
[44] C. Santos, B. Ortiz-Reyes, T. Naranda, M. Remacha, J.P.G. Ballesta, Biochemistry 32 (1993) 4231.
[45] F. Sanchez-Madrid, F. Vidales, J.P.G. Ballesta, Eur. J. Biochem. 114 (1981) 609.
[46] F.J. Vidales, M.T.S. Robles, J.P.G. Ballesta, Biochemistry 23 (1984) 390.
[47] S. Zinker, J.R. Warner, J. Biol. Chem. 251 (1976) 1799.
[48] D. Finley, B. Bartel, A. Varshavsky, Nature 338 (1989) 394.
[49] J.M. Cherry, C. Adler, C. Ball, S.A. Chervitz, S.S. Dwight, E.T. Hester, Y. Jia, G. Juvik, T. Roe, M. Schroeder, S. Weng, D. Botstein, Nucleic Acids Res. 26 (1998) 73.
[50] T.D. Veenstra, S. Martinovic, G.A. Anderson, L. Pasa-Tolic, R.D. Smith, J. Am. Soc. Mass Spectrom. 11 (2000) 78.
[51] S. Martinovic, T.D. Veenstra, G.A. Anderson, L. Pasa-Tolic, R.D. Smith, J. Mass Spectrom. 37 (2002) 99.
[52] R.A. Zubarev, D.M. Horn, E.K. Fridriksson, N.L. Kelleher, N.A. Kruger, M.A. Lewis, B.K. Carpenter, F.W. McLafferty, Anal. Chem. 72 (2000) 563.
[53] P.A. Demirev, J. Ramirez, C. Fenselau, Anal. Chem. 73 (2001) 5725.
[54] W.J. Henzel, T.M. Billeci, J.T. Stults, S.C. Wong, C. Grimley, C. Watanabe, Proc. Natl. Acad. Sci. USA 90 (1993) 5011.
[55] P. James, M. Quadroni, E. Carafoli, G. Gonnet, Biochem. Biophys. Res. Commun. 195 (1993) 58.
[56] M. Mann, P. Hojrup, P. Roepstorff, Biol. Mass Spectrom. 22 (1993) 338.
[57] D.J.C. Pappin, P. Hojrup, A.J. Bleasby, Curr. Biol. 3 (1993) 327.
[58] M. Mann, M. Wilm, Anal. Chem. 66 (1994) 4390.
[59] W.G. Burton, K.D. Nugent, T.K. Slattery, B.R. Summers, L.R. Snyder, J. Chromatogr. 443 (1988) 363.