ELSEVIER

Journal of Chromatography A, 826 (1998) 167–181

JOURNAL OF CHROMATOGRAPHY A

Examination of micro-tip reversed-phase liquid chromatographic extraction of peptide pools for mass spectrometric analysis

Hediye Erdjument-Bromage[a], Mary Lui[a], Lynne Lacomis[a], Anita Grewal[a], Roland S. Annan[b], Dean E. McNulty[c], Steven A. Carr[b], Paul Tempst[a,*]

[a]*Molecular Biology Program, Memorial Sloan-Kettering Cancer Center, New York, NY 10021, USA*
[b]*Department of Physical and Structural Chemistry, SmithKline Beecham Pharmaceuticals, King of Prussia, PA 19406, USA*
[c]*Protein Biochemistry, SmithKline Beecham Pharmaceuticals, King of Prussia, PA 19406, USA*

Received 7 July 1998; received in revised form 24 August 1998; accepted 24 August 1998

Abstract

Mass spectrometry occupies a central position in most current protein identification schemes. So-called 'mass fingerprinting' techniques rely on composite mass patterns of proteolytic fragments, or dissociation products thereof, to query databases. Keys to successful analysis of ever smaller amounts are sensitivity and complete spectral information, both of which depend for a large part on proper sample preparation. Clean-up and concentration of peptide mixtures over eppendorf gel loading tips filled with chromatographic media (i.e. 'micro-tips') are believed to be quite useful in this regard. We have studied quantitative and qualitative aspects of polypeptide extraction using these small manual devices. Optimization of sample volume and additives, micro-tip bed volume, and eluent composition and volume, all contribute to effective recovery (~65–70%, on average). Improper digest conditions can, in fact, lead to far bigger losses, suggesting the need for at least trace amounts of Zwittergent 3-16. Of particular interest is our finding that partial fractionation, obtained by two-step micro-tip elution, generally results in more and better signals during subsequent mass analysis. Thus, by using optimized micro-tips, in combination with adequate sample handling and instrumentation, direct mass spectrometric identification can be routinely and successfully done in any resource facility type setting.

Keywords: Micro-tip liquid chromatography; Mass spectrometry; Sample preparation; Peptides

1. Introduction

Protein identification has turned into a major business for most microchemistry resource facilities. At the same time, and not entirely by coincidence, the technology of choice to carry out such analyses has shifted from traditional chemical sequencing to mass spectrometric-based strategies. Both the older and newly emerging approaches start out with enzymatic fragmentation of the protein; from there on, the methods diverge. Nowadays, peptide masses can quickly and very accurately be determined from non-fractionated mixtures, and at concentrations below 100 fmol per μl, by either matrix-assisted laser-desorption/ionization (MALDI) or electrospray ionization (ESI) mass spectrometry (MS). The resulting information, a 'peptide mass fingerprint',

*Corresponding author. Fax: +1-212-717-3604; E-mail: p-tempst@mskcc.org

0021-9673/98/$ – see front matter
PII: S0021-9673(98)00705-5

enables protein identification against sequence repositories just as well as if bona fide sequence data were used [1–8].

As was the case with the older amino acid analyzers and protein sequencers, unfortunately, the quality of mass spectrometric data is severely compromised by dirty and contaminated samples. In the case of MALDI and infusion-based ESI, signals are highly concentration dependent as well. Both these techniques benefit, therefore, from having the sample available in the smallest possible volume. This issue was first addressed by Wilm and Mann, who used microcolumns filled with reversed-phase (RP) chromatographic material and eluted straight into glass capillary spray needles by centrifugation [9]. Simplified, manually operated versions, using either loose beads, eppendorf gel-loading pipet tips as the 'column' or prepacked 'micro-guard' columns, have subsequently been developed and shown to provide cleanup in addition to concentrating the sample, resulting in improved mass spectrometric data on minute amounts of gel-purified proteins [10–14]. Full optimization of the pipette tip device, or 'microtip', to yield spectra of still better quality, has awaited detailed investigation of several operational variables.

Here, we report on quantitative and qualitative aspects of peptide extraction using these tips, and propose sample handling and chromatographic conditions that should result in generally improved recovery of polypeptides and expanded information content derived from the spectra.

2. Experimental

2.1. Materials

Acetonitrile (MeCN) was obtained from Burdick and Jackson (Muskegon, WI, USA); trifluoroacetic acid (TFA) from Pierce (Rockford, IL, USA); formic acid (FA) and Ponceau S from Fluka (Ronkonkoma, NY, USA); 4-vinylpyridine from Aldrich (Milwaukee, WI, USA); Zwittergent 3-16 (Zwit 3-16) from Calbiochem (San Diego, CA, USA); Amido black from Serva Biochemicals (Paramus, NJ, USA); β-mercaptoethanol, Coomassie Brilliant blue R-250, nitrocellulose and all chemicals/equipment for casting and running polyacrylamide gels, and for electroblotting, were from Bio-Rad (Richmond, CA, USA). All were of the highest possible quality/purity available. All other chemicals and reagents were from J.T. Baker (Philipsburg, NJ, USA), except where noted.

Peptides Pepep-II (ISC_{pe}WAQIGKEPITFEHINYERVSDR), SKI-5 (LQYTEHQQLEGWK), APID (GNNRPVYIPQPRPPHPRL), P8930 (YIPQPRPPHPRL) and Ho1017 (PPRPPHPRL) were chemically synthesized (by Brian Marley and San San Yi at the Sloan-Kettering Microchemistry Facility), *S*-alkylated with 4-vinylpyridine (Pepep-II only), purified by reversed-phase high-performance liquid chromatography (RP-HPLC), quantitated by amino acid composition analysis (at the Rockefeller University Microchemistry facility), diluted to a concentration of 200 pmol per μl water, aliquoted and stored at −70°C, all according to published procedures [15–17]. Pepep-II and SKI-5 (2 nmol stock solution) were iodinated using Iodo-beads (Pierce) and Na ^{125}I (0.4 mCi; Amersham, Arlington Heights, IL, USA) according to the manufacturer's instructions, and separated from unincorporated label by one or two passages over C_{18} Sep-Pak Vac cartridges (Waters, Milford, MA, USA) to yield a solution of labeled peptide with over 99% trichloroacetic acid (TCA) – precipitable counts [checked analytically in the presence of 5 μg of bovine serum albumin (BSA) as carrier]. Counting was done using a LKB Wallac 1272 Clinigamma instrument (Pharmacia, Piscataway, NJ, USA), and the final preparation aliquoted and stored at −70°C.

Yeast glucose-6-phosphate dehydrogenase (G6PD) was purchased from Sigma (St. Louis, MO, USA); a stock solution was made in water and quantitated by amino acid analysis. Selected quantities of this G6PD stock were also diluted in Laemmli sample buffer [18] and subjected to sodium dodecylsulfate–polyacrylamide gel electrophoresis (SDS–PAGE), followed by staining with Coomassie blue and destaining, and excision of the gel pieces. SDS–PAGE-purified protein pMG11 (electrotransferred to nitrocellulose and visualized by Amido black staining) was a kind gift from Dr. Michael Gmachl (Sloan-Kettering Institute, New York, NY, USA). Trypsin was 'modified sequencing grade' from Promega (Madison, WI, USA) and *Achromobacter*

lyticus endoproteinase Lys-C (lysyl endopeptidase, or 'LysC') was obtained from Wako Chemicals (Richmond, VA, USA); proteases were dissolved, stored and handled as described [19].

2.2. Enzymatic proteolysis

Preparative LysC/Tryptic, sequential double-digestion of yeast G6PD was carried out following standard procedures for 'in-solution' proteolysis [19]. Forty nanomoles of the protein was dissolved in 910 μl of 6 *M* guanidinium·HCl (Gu·HCl), heated for 30 min at 50°C, diluted with 1.5 ml water and 0.3 ml of 2 *M* Tris·HCl buffer pH 8.5, and digested with 80 μg of LysC for 5 h at 37°C. The whole digest mixture was then desalted on a C_{18} Sep-Pak Vacuum cartridge, lyophilized, redissolved in 1 ml of 100 m*M* Tris·HCl pH 8.5 and digested to completion with 16 μg trypsin for 2 h at 37°C. An estimated 250 pmol aliquot of the digest mixture was then fractionated by RP-HPLC (see Fig. 4), using a 2.1-mm Vydac 218TP52 C_{18} column (Separations Group, Hesperia, CA, USA) at a flow of 0.1 ml/min and operated essentially as described [16]; peak fractions were manually collected and stored for mass spectrometric identification. Quantitation of the non-fractionated peptide mixture was done by amino acid composition analysis, the concentration adjusted to 10 pmol per μl of 1% TFA, and the final stock solution dispensed in small aliquots and stored at −70°C. Appropriate dilutions of freshly thawed stock were then made, in solvents of choice, immediately before use.

In-situ tryptic digestion of nitrocellulose-bound protein was done in an ammonium hydrogencarbonate–0.5% Zwittergent 3-16 buffer, as described [17] but without *S*-alkylation. In-gel digests were carried out following published procedure [20], with modifications. Briefly, after dehydration with neat MeCN, the gel pieces are dried down under vacuum and reswollen in a minimal volume of 0.2 *M* NH_4HCO_3 buffer, containing 0.1 μg trypsin and 0.1% Zwittergent 3-16; the gel piece is then just about covered with hydrogencarbonate–Zwittergent solution, and the digest allowed to proceed for 2 h at 30°C. After sonication, the supernatant is removed and eventually combined with the peptide solutions resulting from two additional extractions (using 0.1% aqueous FA–MeCN, 40:60), followed by drying in a SpeedVac (Savant, Holbrook, NY, USA) and reconstitution in 0.1% FA just before micro-tip extraction.

2.3. Reversed-phase micro-tips

A slightly modified procedure [10] was used to prepare and operate the miniature size column, or 'micro-tip' (shown in Fig. 1A), for manual RPLC. Initially, a plug is created using an 'ultra-micro tip'

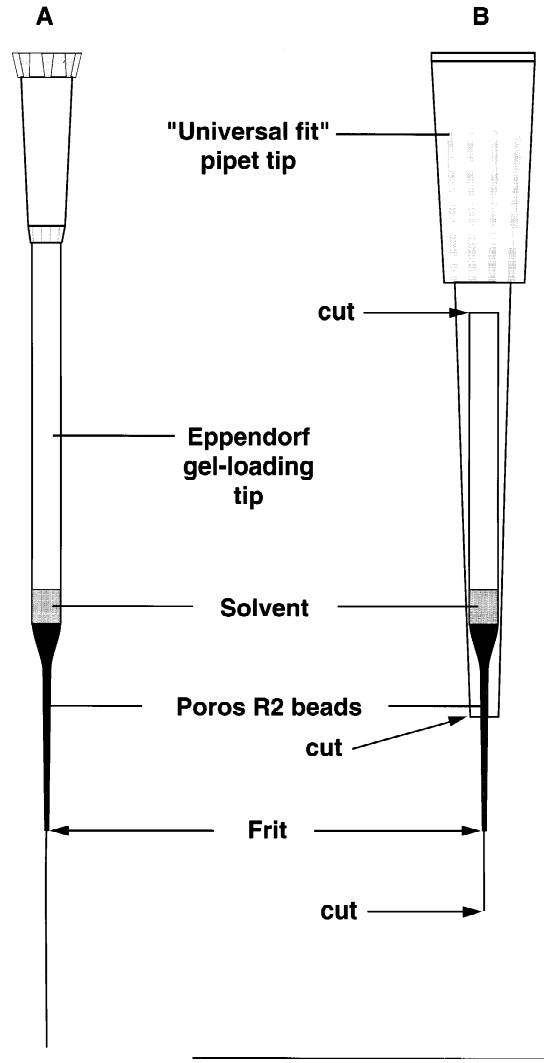

Fig. 1. Schematic diagram of a 'micro-tip'. (A) Micro-tip ready for sample application. (B) Micro-tip, cut with razor blade at top and bottom, inserted in 'universal fit' pipette tip (cut at bottom); ready for elution.

(0.5–10 μl for pipetman; Nortech, Farmingdale, NY, USA) to cork out a small circle from a glass fiber disk (TFA-treated cartridge filter; Cat. No. 400379; PE Applied Biosystems, CA, USA). This is then secured (i.e. pushed down until stuck) into the lower end of an 'Eppendorf gel-loader tip' (Cat. No. 2235165-6; Brinkmann, Westbury, NY, USA) containing 20 μl of absolute ethanol (Florida Distillers, Lake Alfred, FL, USA). Typically, 10 μl of 'sized' (see Section 2.4.) Poros 50 R2 (PerSeptive Bio-Systems, Framingham, MA, USA) bead slurry (4:1; ethanol:bead), which represents a 2-μl packed bead volume, is pipetted on top of the micro-tip frit. The beads are tightly packed under argon pressure (~10–15 p.s.i.), assisted by gentle 'flicking' of the tip, followed by three washes with 20 μl of 0.1% FA (never allowing the beads to dry in between wash steps) (1 p.s.i.=6894.76 Pa). At this point, the micro-tip is checked for steady flow and the absence of dead spaces or air pockets in the column bed (using an illuminated magnifying lens for effective viewing). The final wash is reduced to ~5-mm above bead surface, before loading the sample (restrictions on sample volume, solvent composition and additives may apply; see Section 3 Section 4). Following sample application, the column is washed with three 20-μl volumes of 0.1% FA (again, never forcing the entire wash solution through the beads each time). After the final rinse step, a minimal volume of wash solution is left, just covering the surface of the beads. For optimal elution, the tip is then carefully cut with a razor blade to ~1 cm above the packing (see Fig. 1B); which facilitates subsequent step-wise application of elution solvents. To allow attachment to the argon regulator, however, the tip column must be sleeved by a 'Universal fit pipet tip' (1–200 μl yellow; VWR Scientific, West Chester, PA, USA), cut ~3-mm from the bottom (Fig. 1B). After assembly of this contraption, 3–4 μl of elution solvent (0.1% aqueous FA–MeCN, 84:16) is quickly but carefully pipetted onto the topmost part of the micro-tip, creating an air bubble that separates the wash solvent (still on the beads) from the newly applied elution buffer. The remaining wash solution is then slowly forced out the column without mixing with elution solvent. At the precise moment when the elution buffer touches the surface of the packing, the pressurized tip must immediately be vented and removed from the argon pressure system; 3–4 μl of the second elution buffer (0.1% aqueous FA–MeCN, 70:30) is then applied following the same 'air bubble' routine. Both eluates are collected separately in 0.2-ml micro-tubes (Cat. No. PCR-02; United Scientific Products, San Leandro, CA, USA), stored on ice, and, preferentially, immediately used for mass spectrometric analysis. If long term storage is required, 1 μl of neat MeCN is added to the 16% MeCN-containing eluate.

It should be mentioned that, whereas all the experiments described herein have been done using sized beads, we have since found that 'off-the-shelf' Poros 50 R2 packing works equally well provided a glass frit is inserted at the lower end of the tip (as described above). The frit is essential.

2.4. Sizing of beads

Bead sizing was done by flow flotation [21]. Poros 50 R2 resin (5–25 g) is suspended in toluene in a separatory funnel (500 ml) and it is connected in series to separatory funnels of increasing volumes (1 l, 2 l) by glass tubings making sure that the inlets are almost at the bottom. Toluene is pumped by a peristaltic pump at a very high rate (900 ml/min) through the resin and then through the other separatory funnels in a continuous flow. After several hours beads of decreasing sizes are deposited in larger separatory funnels. The largest beads with measured diameters of 40–60 μm were stored in ethanol and used for tip packing.

2.5. Mass spectrometric analysis

Each peptide pool is usually analyzed twice by MALDI time-of-flight (TOF) MS, in the presence and absence of three peptide calibrants (APID and its truncated forms, P8930 and Ho1017; with respective monoisotopic molecular masses of 2108.155, 1307.762 and 969.575 in the protonated form). Aliquots (0.5 μl) were deposited on the stainless steel probe surface, immediately mixed with an equal volume of α-cyano-4-hydroxycinnamic acid solution (MALDI-Quality premade solution; Bruker-Daltonics, Billerica, MA, USA) on the plate, and allowed to dry at room temperature; calibrants were diluted from concentrated stocks (1 pmol per μl of 0.1% aqueous FA–MeCN, 67:33) and mixed together to yield the required amount of each (typically

6 or 12.5 fmol) per 0.2-μl volume of the same solvent just prior to mixing with the analytes. MALDI-TOF mass spectra were acquired on a REFLEX III (Bruker–Franzen; Bremen, Germany) instrument equipped with a 337-nm nitrogen laser, a gridless pulsed-extraction ion source, SCOUT computerized XY-motion target control and a 2-GHz digitizer; the instrument was always operated in reflector mode (2.9 m total flight path). 25 kV ion acceleration, 26.25 kV reflector and −1.5 kV multiplier voltages were used. Ion extraction was performed 200 ns after each laser irradiance by pulsing down the source extraction lens to 17.5 kV from its initial 25-kV level to give appropriate time-lag focus conditions at the detector. Spectra were obtained by averaging multiple signals; laser irradiance and number of acquisitions (typically 100–150) were operator adjusted to yield maximal peak deflections, derived from the digitizer as TOF data and displayed in real time as mass spectra using a SPARC station 5 (Sun Microsystems; Mountain View, CA, USA). After recalibration with the internal standards, monoisotopic masses were assigned for all prominent peaks (and visually inspected), and a peptide mass list generated.

Electrospray ionization (ESI) MS was done on an API 300 triple quadrupole instrument (PE-SCIEX; Thornhill, Canada), modified with an in*j*ection *a*daptable *f*ine *i*onization *s*ource ('JAFIS') as described in an earlier publication [22]. Needle voltage ranged from 600 to 1350 V depending on the application. The voltages for the orifice and the curtain plate were set at 5 and 350, respectively. Positioning of the source relative to the orifice was accomplished through the instrument control software (ICS) and varied depending on the ionization potential and flow-rate [23]. Q1 scans were collected using a 0.5 u step size, and a 3-ms dwell time over a mass range from 400 to 1400 u; scans were averaged for statistical analysis, and Q1 resolution was set such that the charge state of singly, doubly, and triply charged ions could be ascertained. For operation in the MS–MS mode, Q1 was set to transmit the complete isotopic envelope of the parent. All spectra were averaged with a 0.5 u step size and a 3-ms dwell time for 5 min over the mass range of the singly charged m/z. Q3 resolution was set such that the charge state of the fragment ions could be distinguished. Collision energies, as well as collision associated dissociation (CAD) gas pressures, were optimized individually for each peptide so as to obtain the best MS–MS spectra.

Selected, 'major' mass values (combined from the 16% and 30% peptide pools, but restricted to 900 u$\leq m/z \leq$3000 u) from the MALDI-TOF experiments were arbitrarily taken to search a protein non-redundant database (NRDB; European Bioinformatics Institute, Hinxton, UK) using the PeptideSearch [2] algorithm (PEPTIDESEARCH software and NRDB index file obtained from Dr. M. Mann; EMBL, Heidelberg, Germany). A molecular mass range of up to 300 000 was covered, with a mass accuracy restriction of 30 ppm or better, and a maximum of one missed cleavage site allowed per peptide. After a tentative identification was made, as many as possible of the experimental masses were fitted to the listed sequence (monoisotopic mass values), allowing for maximal 0.1 Da discrepancy. MS–MS spectra from the ESI triple quadrupole analyses were inspected for uninterrupted y″ ion series using the 'find higher AAs' routine of the BioToolbox (PE-SCIEX) software; the resultant information (2–5 amino acids partial sequence, plus corresponding precursor and fragment ion masses) is semi-automatically transferred, by way of a custom AppleScript (Apple Computer, Cupertino, CA, USA), to the SEQUENCETAG [7] program (included in PEPTIDESEARCH software) and used as a search string, with a 2-Da mass error restriction. In case only two y″ ions could be tagged, this limited information was then taken (together with the precursor ion mass) to search the database using the 'Pepfrag' protein identification program from the PROWL resource [24], available over the Internet (http://prowl.rockefeller.edu/PROWL/pepfragch.html). Any protein identification thus obtained was verified by comparing the computer-generated fragment ion series of the predicted tryptic peptide with the experimental MS–MS data; this also allowed to discriminate true from false positives in case more than one protein has been retrieved.

3. Results and discussion

3.1. Solvent selection for elution of micro-tips

Since the use of RP microtips is, in fact, low-

performance liquid chromatography, several operational parameters can be conveniently copied from traditional RP-HPLC [16,25]. Thus, an acetonitrile gradient, or one-step batch-elution, in dilute organic acid is a reasonable first choice for recovery of bound peptides. Because the technique is largely intended for use with sub-pmol amounts of polypeptides, and specifically as a front-end to mass spectrometry, two essential criteria need to be satisfied by the choice of a suitable elution solvent. First, it must prevent analyte losses (to tubes and pipette tips) during transfers; and, secondly, it should result in the highest ionization rates during matrix-assisted laser-desorption or electrospray, and in minimal interference during subsequent mass measurements (time-of-flight determination or quadrupole scans). To address the first issue, a series of simple transfer experiments (by pipetting from one vial to another) were done using 400 fmol of a 25-amino acid long, radiolabeled peptide dissolved in 10 μl of either 0.1% TFA or FA, each containing increasing concentrations (1 to 30%) of MeCN. This particular peptide (Pepep-II; see Section 2.1) elutes relatively late from a standard C_{18} RP-HPLC column and is prone to adsorptive losses [16,26]; it therefore represents a comparative 'worst-case-scenario'. The data shown in Fig. 2 indicate that, whereas reproducibly better recoveries resulted from the use of TFA over formic acid at MeCN concentrations of 20% and below, essentially quantitative transfer was achieved in the presence of 30% MeCN regardless of acid type. It was, therefore, impossible to improve on this result by using higher acid concentrations.

Nonetheless, we wanted to investigate if formic acid concentration might affect, favorably or adversely, efficiencies of ionization and mass measurement in MALDI-TOF and continuous nanoESI (using a 'JAFIS' source) quadrupole MS. Peptide mixtures (G6PD tryptic fragments) were diluted from a concentrated stock to 40 fmol per μl of 30% MeCN, in 0.1 to 10% aqueous formic acid, and analyzed. Signal intensity for five different peptides, ranging in size from nine to 18 amino acids and of varying hydrophobicity (see Table 2), were then measured (Fig. 3 top and bottom). It appeared that FA concentrations of 5% or more offered no advantage in terms of signal enhancement; to the contrary, peak sizes were invariably lower than those derived

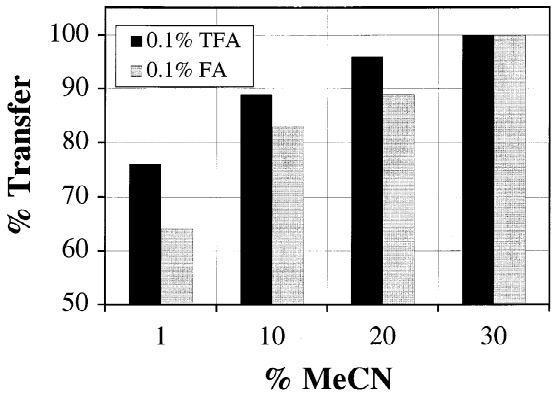

Fig. 2. Peptide losses during transfer between polypropylene tubes. A 40-fmol amount of Pepep-II (sequence under Section 2), spiked with a low femtomole amount of the [^{125}I]-labeled (~20 000 cpm) form, was diluted into 10 μl of either 0.1% trifluoroacetic acid (TFA) or 0.1% formic acid (FA) solutions that also contained increasing concentrations of acetonitrile (MeCN), as indicated. Solutions were then pipetted to a second tube and counted; the emptied tube and pipette tip were also counted. Combined counts of tube A (empty), tube B (transferred solution) and tip were considered 100%. Transfers, presented as % of total, are the result of duplicate experiments.

from identical peptides (at the same concentrations) in MeCN–water solutions containing ≤2% formic acid. Again, 0.1% aqueous FA–MeCN (70:30) proved to be entirely adequate for a variety of peptides tested.

3.2. Packed bed versus sample volume optimization

Reaction volumes of 'in-situ' digestions are generally in the 5- to 20-μl range. We have also shown previously that in the case of 'on-membrane' digestion, the reaction mixtures must be supplemented with 0.5–1% Zwittergent 3-16 to ensure adequate recovery of peptides [17]. This detergent, zwitterionic as the name implies and containing a 16-carbon long tail, binds tightly to RP supports, eluting at considerably higher MeCN concentrations than the large majority of tryptic peptides typically encountered [17]. Thus, in view of the small bed-volumes of RP micro-tips used here, Zwittergent 3-16 could conceivably present a problem in terms of saturating binding sites on the tip. This was experimentally tested by measuring the binding/elution of a radio-

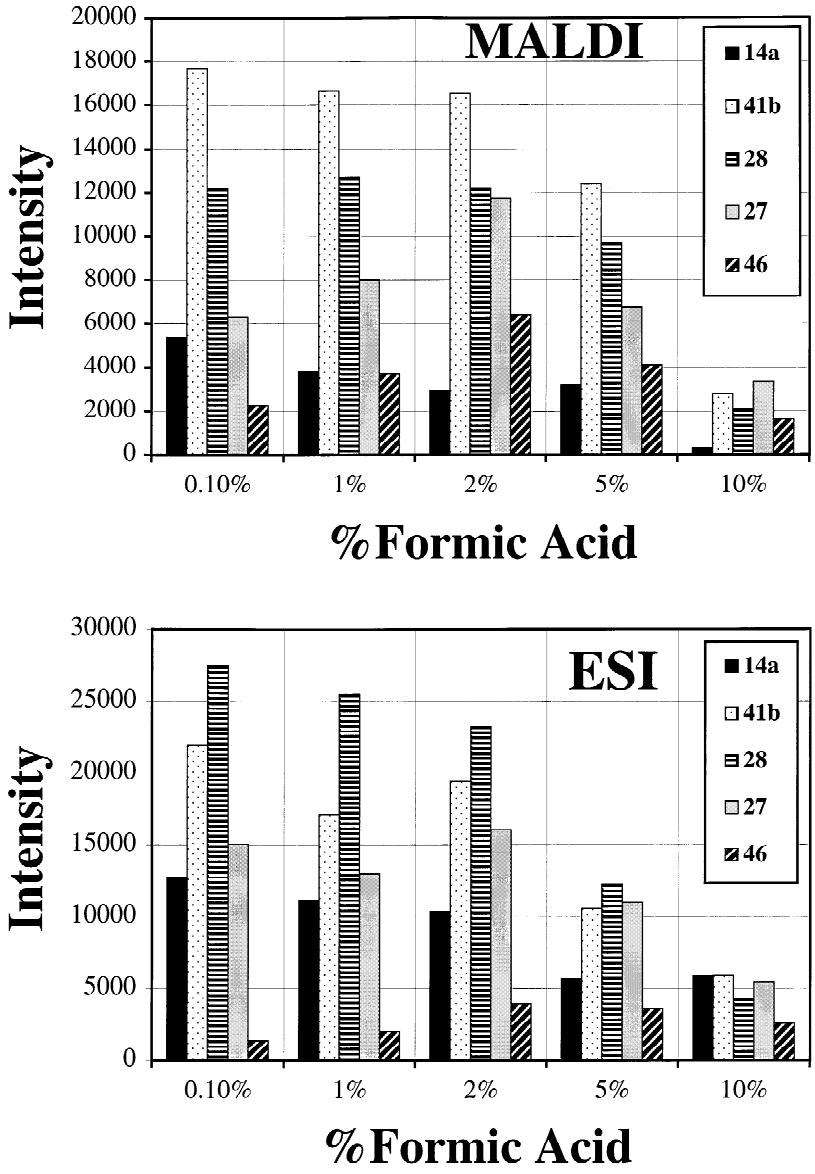

Fig. 3. Effect of formic acid concentration on MS signal strength. Peptide mixtures (G6PD tryptic fragments) were diluted to 40 fmol per μl of 30% MeCN, also containing 0.1 to 10% aqueous formic acid (as indicated), and analyzed by MALDI-TOF-MS ('MALDI') or continuous nano-electrospray quadrupole MS ('ESI') using a JAFIS source [22]. Relative signal intensities for five different peptides are presented. The sequence and monoisotopic masses of each peptide can be found in Table 2 using the key Nos. (#14a, 41b, 28, 27, 46) shown in the inset; corresponding peaks (except No. #46) are also labeled in Fig. 6A and B. A 0.5-μl volume of peptide solution was taken for MALDI-TOF-MS analysis, with signals averaged from 150 acquisitions at constant laser irradiance; results are the mean of two independent experiments. Infusion for ESI-MS analysis was done at 75 nl/min constant flow, and signals averaged from 50 scans (0.5 u step size; 3 ms dwelltime); the same spray needle and back pressure were used for all experiments (done over a 5-h period). More details on both types of mass analysis are to be found in the Section 2.

labeled peptide to/from various sized micro-tips (i.e. bed volumes), using increasing sample volumes and detergent concentrations therein. After loading the sample, the micro-tip was exhaustively washed, and eluted, to eliminate possible 'tailing' effects. The results presented in Table 1 can be summarized as follows. In the presence of 0.5% Zwittergent 3-16, micro-tip bed volume should equal sample volume to yield satisfactory (i.e. two thirds, or better) peptide recoveries; correspondingly, at least a twofold excess (v/v) of beads over sample was required with 1% of the detergent present. Because the absolute amount of Zwittergent is the limiting factor here, increasingly larger volumes can, in fact, be loaded on a micro-tip with fixed bed volume as the concentration of detergent is reduced accordingly (e.g. 10 μl of 0.1% Zwittergent 3-16 on 2 μl of beads; data not shown).

3.3. Stepwise elution from micro-tips

Next, we sought to determine (i) micro-tip elution volumes giving maximal recovery of peptides at the highest possible concentrations, and (ii) whether desorption of any given peptide would occur within a defined, relatively narrow range of organic solvent concentration (i.e. as in actual RP-HPLC). To this end, a radiolabeled peptide of medium hydrophobicity and tightly bound to a RP micro-tip was eluted by sequential addition of aqueous solvents of gradually higher MeCN concentration (from 10 to 30%, in 5% increments). As shown in Fig. 4, adequate recovery was obtained within a window of just 5% and using volumes as small as 1.5–2 times that of the packed bed (as indicated by a virtual absence of 'tailing' during successive elution steps). This result prompted us to explore the possibility of breaking up

Table 1
Peptide binding and recovery to/from RP micro-tips

Bed volume (μl)	Sample volume (μl)	Sample composition					Bound (%)	Eluted (%)
		Zwit 3-16			TFA			
		0.25%	0.5%	1%	1%	10%		
1	10				+		38	36
1	10					+	65	63
1	1			+			43	nd
1	2			+			4	2
1	1		+				74	nd
1	2		+				63	49
1	5		+				33	30
1	5	+					71	65
2	1			+			70	nd
2	2			+			15	12
2	1		+				79	73
2	2		+				74.9±3.6[a]	67.7±3.1[a]
2	5		+				58	nd
2	10		+				24	nd
2	5		+		+		71	nd
2	5		+			+	76	nd

[a] Average of six experiments (±standard deviation).

Note. 1-pmol amounts of [^{125}I]-labeled (~50 000 cpm) Pepep-II (sequence under Section 2) were mixed with 1 pmol of G6PD tryptic digest mixture, diluted in the solvent and total volume as indicated, loaded on a micro-tip containing a measured bed volume of sized Poros 50 R2 beads (see Section 2) and washed three times with 20 μl of 0.1% formic acid (FA). The entire micro-tip was then either directly used for γ-counting ('Bound', expressed as % of total loaded), or eluted with three column volumes of 0.1% aqueous FA–MeCN (60:40) and both eluate ('eluted', expressed as a percentage of total loaded) and the 'used' micro-tip counted (counts eluted and those leftover on the tip were added to give the amount that had been bound). Experiments were done twice and the results were in comparatively good agreement; only one set is listed, except where indicated ([a]), in which case the numbers represent the mean of six experiments (±standard deviation).
nd, not determined.

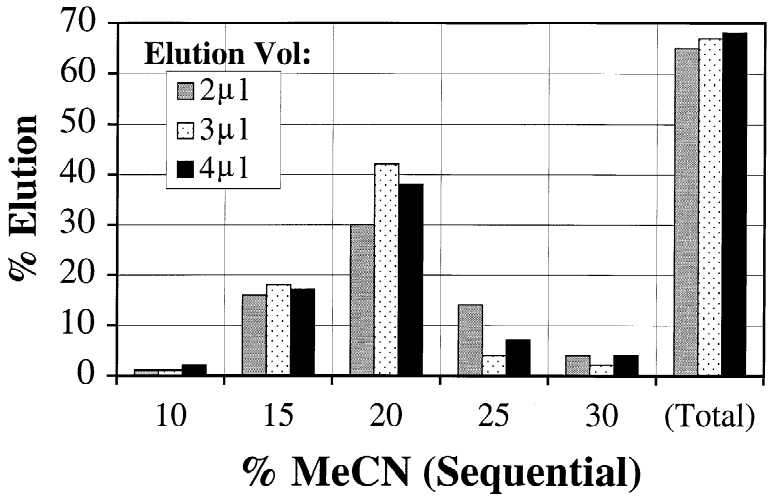

Fig. 4. Sequential elution of peptide from RP micro-tips. One-picomole of SKI-5 (sequence under Section 2), spiked with a low femtomole amount of the [^{125}I]-labeled (~50 000 cpm) form, was diluted in 2 μl of 0.5% Zwittergent 3-16 (in 100 m*M* hydrogencarbonate buffer), loaded on a micro-tip containing a 2-μl bed volume of sized Poros 50 R2 beads (see 'Experimental'), washed three times with 20 μl 0.1% formic acid (FA), and step-wise eluted with constant volumes (2, 3 or 4 μl, in separate experiments) of 0.1% FA, containing increasing concentrations of MeCN (as indicated). Eluates were counted and recovery at each step are shown as the percentage of total peptide loaded; recovery combined from all sequential elution steps is also shown ('Total').

complex peptide mixtures into smaller pools by manual micro-tip chromatography, and how such partial fractionation might benefit subsequent mass spectrometric analysis in terms of total information content.

A tryptic digest of yeast G6PD (M_r~55 000) was used as a model system. Because this substrate is protease resistant in its native form [27], a sequential double-digest (in solution) had to be done; first with endoproteinase LysC in the presence of Gu·HCl, then with trypsin under non-denaturing conditions. The resulting peptide mixture was carefully analyzed by micro-preparative RP-HPLC with MALDI-TOF-MS analysis of the collected peaks (Fig. 5 Table 2). In this way, a total of 53 different peptides (sequences listed in Table 2) were found to be present in the G6PD digest mixture. Although some smaller and bigger (and/or very hydrophobic) peptides may have been overlooked this way, we felt that, rather than by predicting all theoretical peptides, combining bona fide HPLC with MALDI-TOF-MS analysis of purified fractions would yield the maximum number of peptides that could possibly be detected after micro-tip extraction by mass analysis of the various sized pools; it therefore represents a 'practical 100%'

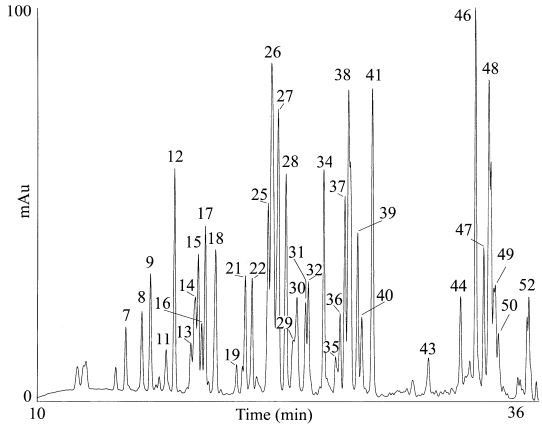

Fig. 5. RP-HPLC separation of G6PD endo-Lys C/tryptic, sequential double-digest. Profile of an estimated 250-pmol digest (performed as described under Section 2) separated on a 2.1-mm Vydac 218TP52 C_{18} column using a linear two-step acetonitrile gradient (in 0.1% TFA) of 3.5–35%/22.5 min, 35–70%/12.5 min at a flow-rate of 0.1 ml/min and at ambient temperature. Full scale corresponds to 0.1 AU at 214 nm; the time scale is from 10 to 38 min. Peak fractions were manually collected and analyzed by MALDI-TOF-MS; the numbering of the peaks corresponds to the listing in Table 2.

Table 2
Peptides generated by endo-LysC/tryptic sequential double-digest of yeast G6PD

HPLC no.	Sequence	Position	MH+ (mono)	ESI	MALDI 1 Step	MALDI Multi
7	EYMQK	476–480	698.319			
8	FGTEGR	226–231	666.321			
9a	LQYK	353–356	551.320			
9b	ELVK	191–195	517.299			
11	VEIR	349–352	516.315			
12a	ERFGTEGR	224–231	951.465			+
12b	SANVDVPHR	113–121	994.507	+	++	++
13a	LSMEEDLKSR	54–63	1223.594			
13b	SEDGSKPAYVDDDTVDKDSK	296–315	2170.963			
14a	VYAENGITR	143–151	1022.527	+	+	++
14b	SEDGSKPAYVDDDTVDK	296–312	1840.809	+		
15	YVMPEK	483–488	766.381			+
16	EEELYR	178–183	838.395			+
17	IDHYLGK	184–190	845.452			+
18	EGYLDPSTK	37–45	1009.484	+		
19	ac-SEGPVKFEK	1–9	1063.543			
21a	AVASGVFK	357-364	778.447			
21b	PFGHDLASAR	157-166	1070.538		+	++
22	VLPHLK	64–69	706.462			+
25	IFGYAR	46–51	726.294		++	++
26a	VQPDAAVYLK	375–384	1103.610	++	+	+
26b	HPYAWPVTKPEDTK	489–502	1668.839	++	+	++
27	HIERPDGPTPEIYPYGSR	452–469	2084.020	+	++	++
28	VIVEKPFGHDLASAR	152–166	1638.897	++	++	++
29	DIPNNELVIR	365–374	1182.648	++	++	++
30a	DALLGDHSNFVR	424–435	1343.671	+	++	++
30b	NTVISVFGASGDLAKKK	10–26	1734.975	++	+	++
31	KPHGEADDSKVEQFFK	70–85	1861.909			+
32	DNIQSVQISFK	213–223	1278.670	+	+	+
34	TPGLSNATQVTDLNLTYASR	389–408	2122.078	++	++	++
35	NLLVLR	195–200	727.483			
36	NLGPLFK	171–177	788.467			+
37	MVSYISGNYDTDEGFDELR	86–104	2210.917	++	++	++
38a	FGNQFLNASWNR	201–212	1453.698	++	++	++
38b	AVAPIDTDDVLLGQYGK	279–295	1774.923	++	+	++
39	NTVISVFGASGDLAK	10–24	1478.786	++	+	+
40	NLGPLFKEEELYR	171–183	1607.844		+	++
41a	WEGVPIMMR	331–339	1118.549	+	+	++
41b	GGYFDSIGIIR	232–242	1197.627	++	++	++
43	WEGVPIMMRAGK	331–342	1374.703			
44a	KTFPALFGLFR	26–36	1296.747			
44b	EYMQKHKYVMPEK	476–488	1710.835			
44c	NLLVLRFGNQFLNASWNRDNIQSVQUSFK	195–223	3421.814			
46	YQDFWIPEAYEVLIR	498–423	1941.975			
47	TPGLSNATQVTDLNLYASRYQDFWIPEAYEVLIR	389–423	4045.034			
48a	TFPALFGLFR	27–36	1168.652			+
48b	DALLGDHSNFVRDDELDISWGIFTPLLK	424–451	3186.611			
48c	DVMQNHLLQIMTLLTMERPVSFDPESIRDEK	243–273	3685.839			
49a	LFYLALPPSVFLTVAK	122–137	1779.046			
49b	DDELDISWGIFTPLLK	436–451	1861.959			+
50	DVMQNHLLQIMTLLTMERPVSFDPESIR	243–270	3313.675			
52a	GGYFDSIGIIRDVMQNHLLQIMTLLTMERPVSFDPESIRDEK	232–273	4364.447			
52b	TPGLSNATQVTDLNLTYASRYQDFWIPEAYEVLIRDALLGDHS NFVRDDELDISWGIFTPLLK	489–451	7212.627			

Note. All peptides recovered from a semi-preparative RP-HPLC separation, and identified by MALDI-TOF-MS analysis of highly purified fractions, are listed; numbers (under 'HPLC no.') correspond to labeled peaks on the chromatogram shown in Fig. 5. Sequence, position in the G6PD sequence, and calculated monoisotopic mass (protonated form; 'MH+') are also given for each peptide. Peptides for which major (++) or medium-sized (+) signals were observed during continuous flow electrospray (JAFIS source) quadrupole scan or MALDI-TOF-MS of digest mixtures that had been extracted over micro-tips with single-step elution (spectra shown in Fig. 6A and B) are marked in the 'ESI' and 'MALDI 1 step' columns; those observed in the combined MALDI-TOF spectra of multi-step eluted peptides (shown in the various panels of Fig. 7) are tagged under 'MALDI Multi'.

recovery scenario. G6PD tryptic peptides (200 fmol) were thus adsorbed to an appropriately sized micro-tip and eluted, either in a single step (4 μl of 0.1% aqueous FA–MeCN, 70:30) or stepwise (10–30% MeCN in four equal increments), and the eluates analyzed by MALDI-TOF-MS (Fig. 6A Fig. 7) and JAFIS(ESI)-MS (Fig. 6B); peak signals were then matched to the available list (Table 2). Single-step eluates resulted in largely overlapping peak patterns as obtained by the two different types of MS analysis; each about 20 signals over background, half of them major (arbitrary assignment). The majority of peptides thus identified mapped to the center portion of the HPLC trace shown in Fig. 5 (from peak 21 to 41); the thirty or so peptides (or, better, MS signals representing those peptides) apparently missing from the spectra were either not recovered from the tip or MALDI and ESI ionizations must have been suppressed [28]. On the other hand, multi-step elution of a similar micro-tip yielded a combined set MALDI-TOF-MS signals (Fig. 7) exceeding those of the one-step elutions both in number (29 peaks) and overall quality (17 major signals). As illustrated by the spectra and shown in Table 2, the extra MS signals are derived from either small, hydrophilic peptides or very hydrophobic ones, corresponding to early or, respectively, late positions in a regular HPLC chromatogram. Multi-step elution may therefore offer a significant advantage by virtue of (i) additional signals when analyzing digests of either large proteins or simple mixtures (two to three proteins), or (ii) detecting unusually hydrophilic (e.g. phosphorylated) or hydrophobic polypeptides. In its most simple form, we determined that a two-step elution is best done with MeCN concentrations of 16% and 30% (data not shown). Yet, even under those optimized conditions, there will always be some peptides missing.

3.4. Practical applications

Micro-tip peptide pool extraction was then tested using a 'real-life' type sample, namely 'in-gel' tryptic digestion of 100 fmol G6PD (amount loaded on the gel). Although not the scope of the current study, we decided to introduce an extra variable here by supplementing the standard bicarbonate digest buffer (~10 μl) in one experiment with 0.1% Zwittergent 3-16. The rationale for this experiment was our earlier observation that the detergent, in addition to dissociating polypeptides from membranes, also prevents adsorptive losses to glass and plastic [17]. At the end of the procedure (Section 2), the combined extraction volumes were dried down, including Zwittergent (a powder) when present, redissolved in 10 μl of 0.1% formic acid (with brief warming at 37°C to solubilize the detergent), the entire volume loaded on a micro-tip (2-μl bed volume) and eluted in a single step. MALDI-TOF-MS analysis of the eluates, pictured in Fig. 8, fully illustrates the two important conclusions that could be drawn from these experiments. Firstly, digests done in the presence of reduced Zwittergent concentrations yielded good quality, easily interpreted MALDI-TOF spectra (Fig. 8; top panel); from a total of 20 peaks, 14 are derived from G6PD peptides and three from trypsin autolytic products (the three others are of undefined nature). The absence of detergent from the digest mixtures, on the other hand, resulted in spectra nearly devoid of any useful information; only a few small peaks could be produced after prolonged analysis (Fig. 8; bottom panel). It thus appeared that, under ideal conditions, peptides derived from as little as 100 fmol of protein prepared by SDS–PAGE can be satisfactorily analyzed by MS after RP micro-tip extraction. The importance of sample preparation, however, cannot be sufficiently emphasized.

The analytical procedure as outlined above has subsequently been validated by the results of numerous identification experiments of gel-purified proteins at the sub-pmol level. Mass spectral data of one such application is presented in Fig. 9. The protein (M_r~24 000) was submitted for analysis in membrane-bound form, digested in the presence of 0.5% Zwittergent, and one half loaded on a micro-tip and eluted in two steps (16% and 30% MeCN in 0.1% aqueous FA). Six major m/z peak signals, combined from two MALDI-TOF analyses (and annotated in Fig. 9A and B), were arbitrarily selected and taken to search a non-redundant database using the 'PeptideSearch' algorithm [2]. Five out of the six m/z values matched monoisotopic masses of predicted tryptic fragments of a single protein (represented twice in the database), the endoplasmatic reticulum protein 'ERS-24' (GenBank U91742) [29]; the sixth

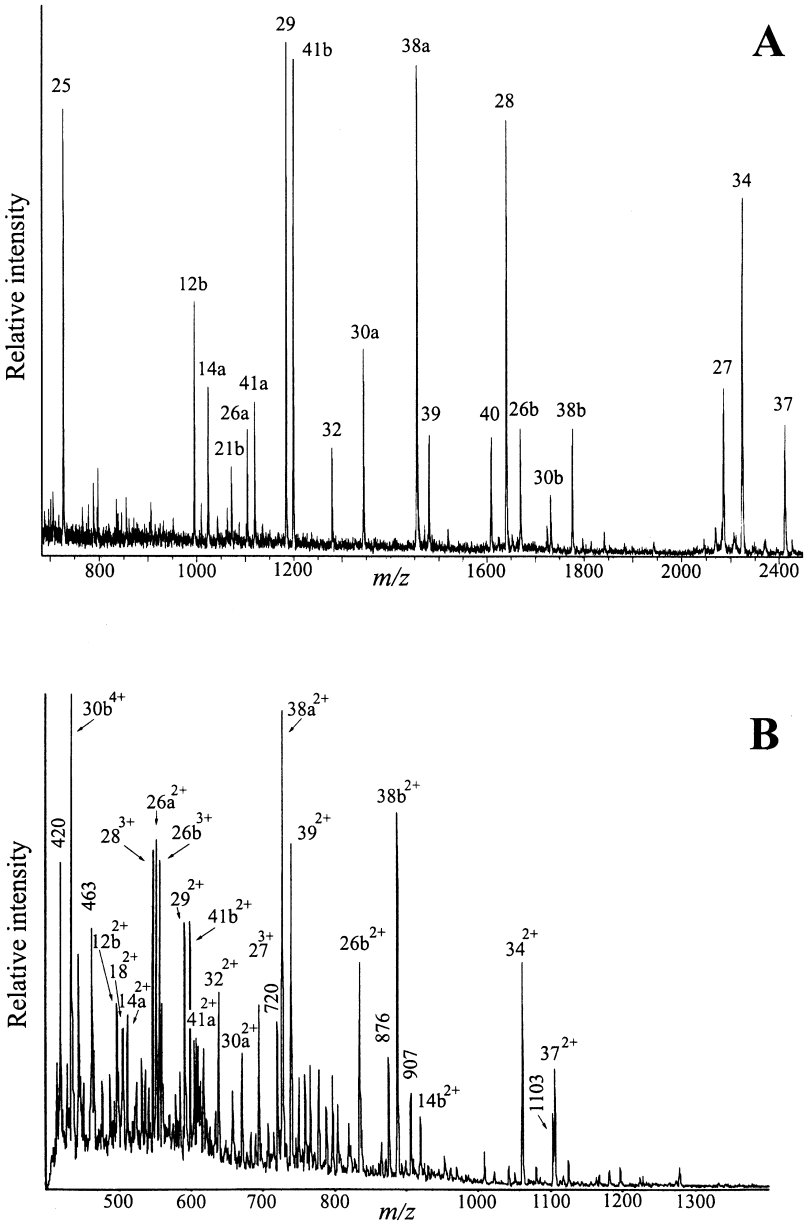

Fig. 6. Mass spectra of G6PD tryptic peptides after one-step RP micro-tip extraction. A 200-fmol amount of G6PD tryptic digest (prepared as described under Section 2) was diluted in 2 μl of 0.5% Zwittergent 3-16 (in 100 m*M* hydrogencarbonate buffer), loaded on a micro-tip containing a 2-μl bed volume of sized Poros 50 R2 beads, washed three times with 20 μl of 0.1% formic acid (FA), and eluted with 4 μl of 0.1% aqueous FA–MeCn (70:30). Eluates were then taken for (A) MALDI-reflectron (re)-TOF (0.5 μl deposited) and (B) continuous flow nanoESI (JAFIS source operated at 75 nl/min) quadrupole scan analysis (see Section 2 Fig. 3). Numbering of the peaks corresponds to the listing in Table 2; peptide charge states are indicated on the JAFIS-Q1 spectrum as superscripts (panel B).

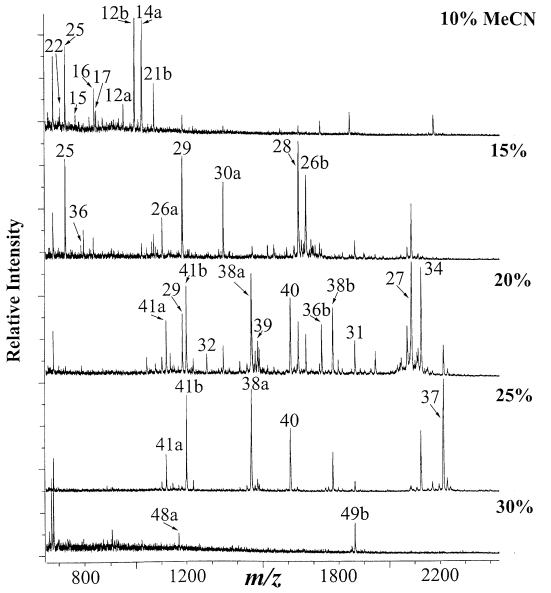

Fig. 7. MALDI-TOF mass spectra of G6PD tryptic peptides recovered from multi-step RP micro-tip extraction. A 200-fmol amount of G6PD tryptic digest was prepared and used for micro-tip extraction as described under Fig. 6, except that step-wise elution was done with 4 μl of 0.1% FA, containing increasing concentrations of acetonitrile (10–30% MeCN), as indicated on the panels. All five eluates were then taken for MALDI-reTOF analysis (0.5 μl deposited), which was done under the exact same conditions as used to generate the spectrum shown in Fig. 6A. Numbering of the major peaks corresponds to the listing in Table 2; peak sizes in the different panels (and those in Fig. 6A) are to scale.

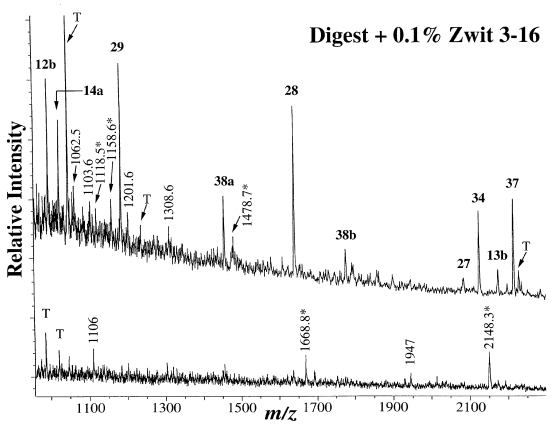

Fig. 8. MALDI-TOF mass analysis of 100 fmols of yeast G6PD, trypsinized in gel and extracted over an RP micro-tip. A 100-fmol amount of G6PD (loaded on the gel) was digested with trypsin as per the procedure described under Section 2, in the presence (top panel) or absence (bottom panel) of 0.1% Zwittergent 3-16, extracted over an RP micro-tip and analyzed by MALDI-reTOF-MS (for more details, see Section 2 and elsewhere in the text). Peaks labeled with bold numbers correspond to peptides listed in Table 2 (using the same numbering), those labeled with 'T' are trypsin autolytic products, those labeled with an m/z value followed by the * mark are predicted G6PD-derived peptides not listed in Table 2 (and thus not generated in a solution digest), all others correspond to molecules of undefined nature.

m/z value corresponded to an ERS-24 peptide containing oxidized Met. By contrast, random hits with various database entries, at 30 ppm error tolerance, were obtained only when a query was done for any three matches out of all six experimental masses. Analysis of the '30% pool' by JAFIS-(ESI)–MS–MS (Fig. 9C and D) allowed independent confirmation of this identification by searching the database, using a 'SequenceTag' as explained under Fig. 9.

4. Conclusions

The need for application laboratories to continually identify, and optimize, all experimental variables of MS-based protein identification prompted the current study of one such critical step, peptide extraction using RP micro-tips. Since we were mostly concerned with achieving a high level of confidence, special emphasis was placed on improving conditions to result in spectra with expanded and better quality information content; this should include signals on short hydrophilic and long hydrophobic peptides as well. In addition, the ideal method must tolerate sizable amounts of gel-pieces or membranes but only generate the lowest background (e.g. protease autolytic fragments, environmental and chemical noise), all the while achieving the highest possible signal-to-noise ratios. Here, we describe a thorough optimization of an existing, manual RP microextraction technique [10]. We find that, when used in conjunction with adequate instrumentation and sample handling procedures, MALDI-TOF and continuous nanoESI (JAFIS)-based identification can be done with as little as 100 fmol of the protein in the gel.

Two issues proved to be particularly important for successful analysis at those levels of sensitivity; the

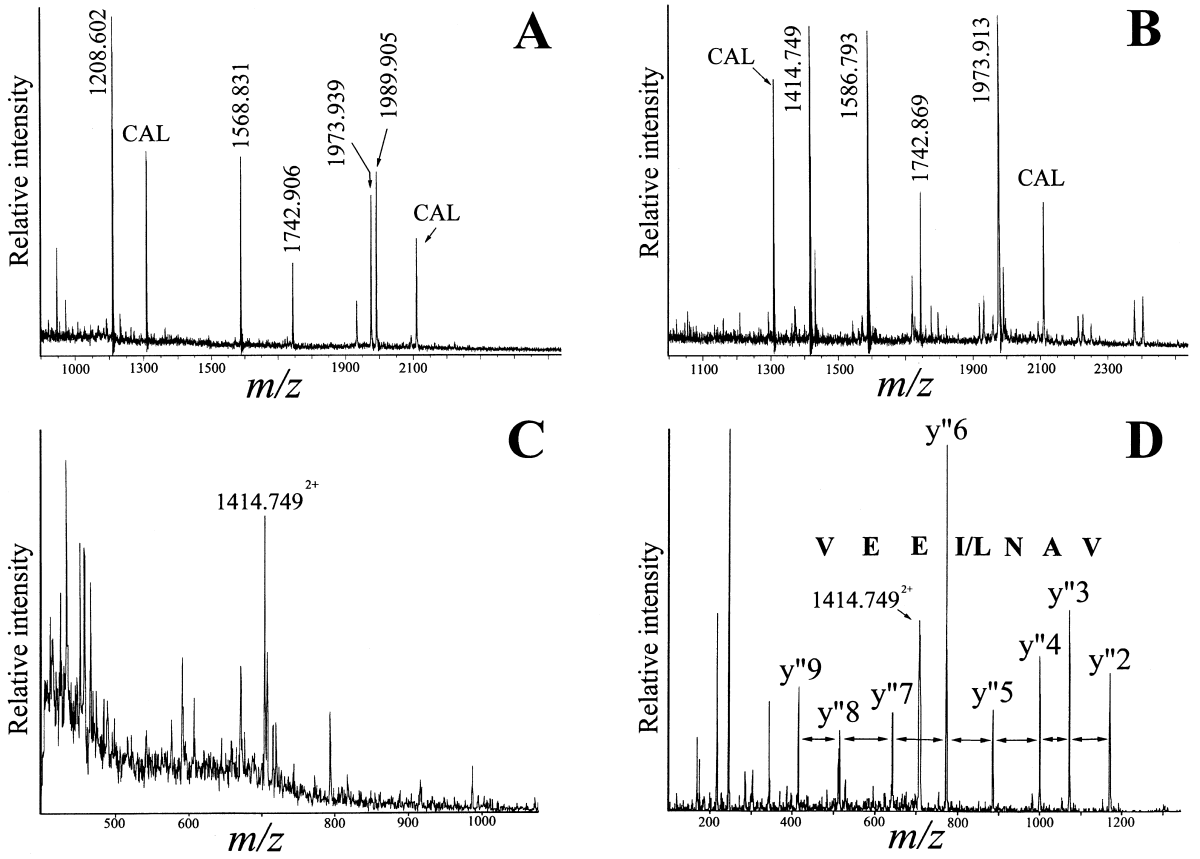

Fig. 9. Identification of a M_r 24 000 vesicle trafficking protein by mass spectrometry. The tryptic digest mixture was passed over an RP micro-tip and the peptides batch fractionated into a 16% and 30% pool. Each pool was individually analyzed by MALDI-reTOF-MS (16% fraction shown in panel A; 30% fraction in panel B) and by continuous flow ESI (JAFIS) triple quadrupole MS–MS (Q1 scan of 30% fraction shown in panel C); only the relevant portions of the spectra are shown. Both types of MS analysis served to independently identify this M_r 24 000 protein as the endoplasmatic reticulum v-SNARE, 'ERS-24' (GenBank U91742) [29]. MALDI-reTOF mass spectra were obtained by averaging 150 scans under constant irradiance. The calibrants (CAL) and the six most prominent peaks (from both pools combined) are labeled in panels A and B; the corresponding m/z values were taken to query a non-redundant protein sequence database (NRDB) for pattern matches, using the PEPTIDESEARCH program [2]. With a requirement of five matches out of six, at a mass accuracy of 30 ppm or better, and a maximum of one missed cleavage site per peptide, a single protein was retrieved (entered in the database twice). The ESI-MS (Q1) spectrum of the 30% fraction (obtained by a JAFIS-generated continuous flow of 4 nl/min, and averaging 100 scans; panel C) contained several peaks corresponding to those observed by MALDI-TOF-MS analysis of the same pool (panel B). One peptide (1414.749^{2+}; panel C) was then selected, by appropriate tuning of Q1, for collision-induced dissociation and subsequent analysis of fragment spectra (in Q3), as shown in panel D. A short sequence was assigned, based on the presence of a contiguous y″ ion series, enabling positive identification of the protein by SequenceTag (peptide M_r 1414±1; [416.0]VEE[IL]NAV[1170.5]) based searching of the NRDB database [7].

partial fractionation of complex mixtures before mass analysis and the effective recovery at all steps. The former is most easily accomplished by a manual, two-step elution of the micro-tip (at 16% and 30% MeCN). Recoveries can be adversely affected at the stages of sample preparation (including proteolysis/transfer to the micro-tip), manual extraction itself, and introduction into mass spectrometers. We show that about two thirds of an average peptide is recovered from the micro-tip. Improper digest conditions, on the other hand, can lead to far bigger losses, as clearly illustrated by some of the data presented here; suggesting the need for at least trace amounts of Zwittergent 3-16. Furthermore, eluting

micro-tips in unnecessarily high concentrations of formic acid will likely produce adverse affects on ionization during mass analysis. Rigorous adherence to sample handling protocols is thus on order when working at the highest levels of sensitivity.

Acknowledgements

The authors are indebted to San San Yi, Natalya Buga and Brian Marley for peptide synthesis, Scott Geromanos and Gordon Freckleton for ESI-MS–MS analysis, Tim Bromage for help with the figures, Matthias Mann for providing PEPTIDESEARCH software, and Bill Henzel for useful suggestions on micro-tip packing. This work was supported by Development Funds from NCI Grant P30 CA08748.

References

[1] W.J. Henzel, T.M. Billeci, J.T. Stults, S.C. Wong, C. Grimley, C. Watanabe, Proc. Natl. Acad. Sci. USA 90 (1993) 5011–5015.
[2] M. Mann, P. Højrup, P. Roepstorff, Biol. Mass Spectrom. 22 (1993) 338–345.
[3] D.J.C. Pappin, P. Højrup, A.J. Bleasby, Curr. Biol. 3 (1993) 327–332.
[4] J.R. Yates III, S. Speicher, P.R. Griffin, T. Hunkapiller, Anal. Biochem. 214 (1993) 397–408.
[5] P. James, M. Quadroni, E. Carafoli, G. Gonnet, Protein Sci. 3 (1994) 1347–1350.
[6] S.D. Patterson, R. Aebersold, Electrophoresis 16 (1995) 1791–1814.
[7] M. Mann, M. Wilm, Anal. Chem. 66 (1994) 4390–4399.
[8] J.K. Eng, A.L. McCormack, J.R. Yates III, J. Am. Soc. Mass Spectrom. 5 (1994) 976–989.
[9] M. Wilm, M. Mann, Anal. Chem. 68 (1996) 1–8.
[10] R.S. Annan, D. McNulty, M.J. Huddleston, S.A. Carr, International Symposium ABRF'96: Biomolecular Techniques, San Francisco, CA, 1996, Abstract S73.
[11] H. Zhang, P.E. Andren, R.M. Caprioli, J. Mass Spectrom. 30 (1995) 1768–1771.
[12] P.L. Courchesne, S.D. Patterson, BioTechniques 22 (1997) 246–250.
[13] K. Gevaert, H. Demol, M. Puype, D. Broeckaert, S. De Boeck, T. Houthaeve, J. Vandekerckhove, Electrophoresis 18 (1997) 2950–2960.
[14] F. Rusconi, J.-M. Schmitter, J. Rossier, M. le Maire, Anal. Chem. 70 (1998) 3046–3052.
[15] P. Casteels, J. Romagnolo, M. Castle, K. Casteels-Josson, H. Erdjument-Bromage, P. Tempst, J. Biol. Chem. 269 (1994) 26107–26115.
[16] C. Elicone, M. Lui, S. Geromanos, H. Erdjument-Bromage, P. Tempst, J. Chromatogr. A 676 (1994) 121–137.
[17] M. Lui, P. Tempst, H. Erdjument-Bromage, Anal. Biochem. 241 (1996) 156–166.
[18] U.K. Laemmli, Nature 227 (1970) 680–685.
[19] L.R. Riviere, P. Tempst, in: J.E. Coligan, B.M. Dunn, H.L. Ploegh, D.W. Speicher, P.T. Wingfield (Eds.), Current Protocols in Protein Science, Greene Publishing Associates and Wiley Interscience, Brooklyn, NY, 1995, Unit 11.1., pp. 1–19.
[20] U. Hellman, C. Wernstedt, J. Gonez, C.-H. Heldin, Anal. Biochem. 224 (1995) 451–455.
[21] B. Chenera, J.A. Finkelstein, D.F. Veber, J. Am. Chem. Soc. 117 (1995) 11999–12000.
[22] S. Geromanos, J. Philip, G. Freckleton, P. Tempst, Rapid Commun. Mass Spectrom. 12 (1998) 551–556.
[23] S. Geromanos, P. Tempst, unpublished results.
[24] D. Fenyo, W. Zhang, B.T. Chait, R.C. Beavis, Anal. Chem. 68 (1996) 721A–726A.
[25] P. Tempst, M.W. Hunkapiller, L.E. Hood, Anal. Biochem. 137 (1984) 188–195.
[26] H. Erdjument-Bromage, S. Geromanos, A. Chodera, P. Tempst, in: R.H. Angeletti (Ed.), Techniques in Protein Chemistry IV, Academic Press, San Diego, CA, 1993, pp. 419–426.
[27] H. Erdjument-Bromage, P. Tempst, unpublished results.
[28] T.M. Billeci, J.T. Stults, Anal. Chem. 65 (1993) 1709–1716.
[29] I. Paek, L. Orci, M. Ravazzola, H. Erdjument-Bromage, M. Amherdt, P. Tempst, T.H. Sollner, J.E. Rothman, J. Cell Biol. 137 (1997) 1017–1028.