Available online at www.sciencedirect.com
ScienceDirect
International Journal of Mass Spectrometry 268 (2007) 147–157
International Journal of Mass Spectrometry
www.elsevier.com/locate/ijms

Solid phase extraction-liquid chromatography (SPE-LC) interface for automated peptide separation and identification by tandem mass spectrometry

Ole Bjeld Hørning [a,b], Søren Theodorsen [b], Ole Vorm [b], Ole Nørregaard Jensen [a,*]

[a] *Department of Biochemistry and Molecular Biology, University of Southern Denmark, Campusvej 55, DK-5230 Odense M, Denmark*
[b] *Proxeon Biosystems A/S, Staermosegaardsvej 6, DK-5230 Odense M, Denmark*

Received 18 March 2007; received in revised form 21 June 2007; accepted 21 June 2007
Available online 28 June 2007

Dedication to Peter Roepstorff in celebration of his 65th birthday.

Abstract

Reversed-phase solid phase extraction (SPE) is a simple and widely used technique for desalting and concentration of peptide and protein samples prior to mass spectrometry analysis. Often, SPE sample preparation is done manually and the samples eluted, dried and reconstituted into 96-well titer plates for subsequent LC–MS/MS analysis. To reduce the number of sample handling stages and increase throughput, we developed a robotic system to interface off-line SPE to LC–ESI-MS/MS. Samples were manually loaded onto disposable SPE tips that subsequently were connected in-line with a capillary chromatography column. Peptides were recovered from the SPE column and separated on the RP-LC column using isocratic elution conditions and analysed by electrospray tandem mass spectrometry. Peptide mixtures eluted within approximately 5 min, with individual peptide peak resolution of ∼7 s (FWHM), making the SPE-LC suited for analysis of medium complex samples (3–12 protein components). For optimum performance, the isocratic flow rate was reduced to 30 nL/min, producing nanoelectrospray like conditions which ensure high ionisation efficiency and sensitivity. Using a modified autosampler for mounting and disposing of the SPE tips, the SPE-LC–MS/MS system could analyse six samples per hour, and up to 192 SPE tips in one batch. The relatively high sample throughput, medium separation power and high sensitivity makes the automated SPE-LC–MS/MS setup attractive for proteomics experiments as demonstrated by the identification of the components of simple protein mixtures and of proteins recovered from 2DE gels.

Keywords: SPE; LC; ESI-MS/MS

1. Introduction

Mass spectrometry (MS) is now established as the method of choice for comprehensive analysis of protein and peptide samples in proteomics research [1,2]. Mass spectrometry is usually combined with biochemical, genetic or immunological methods for protein sample preparation and separation. For "bottom-up" proteomics, the protein sample is then digested with enzymes to generate peptides, which are amendable to mass spectrometry analysis. Simple samples composed of only few protein components generate on the order of 50–100 peptides which are suited for static electrospray ionisation (ESI)-MS or matrix assisted laser desorption/ionisation (MALDI)-MS analysis. Complex samples consisting of hundreds of proteins generate thousands of peptides that must be fractionated prior to mass spectrometry analysis, e.g., by multidimensional LC [3,4]. Using SDS–PAGE [5] or 2DE [6] for protein separation after immunoprecipitation or affinity enrichment often generates protein samples that consist of 3–15 components. In this case neither the direct nanoelectrospray nor the LC based methods are particularly well suited inasmuch as direct MS/MS does not provide sufficient data for a comprehensive analysis, whereas the LC–MS/MS analysis is too time consuming. Thus, there is a need for efficient analytical methods for characterization of such medium-complex protein samples.

Several technological solutions have been engineered to fill out the gap between the simple methods using either MALDI-MS/MS or static nanoelectrospray MS/MS and the

* Corresponding author. Tel.: +45 65 50 23 68.
E-mail address: jenseno@bmb.sdu.dk (O.N. Jensen).

1387-3806/$ – see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijms.2007.06.017

more advanced technique of LC–MS/MS and multidimensional LC–MS/MS. Xie et al. demonstrated separation of simple samples composed of either five intact proteins or five peptides within minutes using gradient elution on a monolithic column [7]. Using a ultrahigh pressure HPLC system (20 kpsi) and 1–10 min gradients coupled to a linear ion trap tandem mass spectrometer, Shen et al. were able to identify 250 proteins from a *S. oneidensis* tryptic digest [8]. Isocratic elution is a technologically simpler alternative to gradient elution. It does not provide the same degree of separation, however, as the solvent composition remains constant throughout the separation, faster elution times can be achieved and the column regeneration time is reduced. A number of procedures for rapid peptide separation by reversed-phase LC using isocratic elution have been used for the analysis of protein samples isolated by SDS–PAGE and 2DE. Chen et al. used high flow rate isocratic elution (2.7 μL/min) until the peptides were on the brink of leaving the column, then the flow rate was lowered to 0.4 μL/min for the MS analysis [9]. Another way of increasing throughput is by using a dual parallel column system. Here two analytical columns are operated by one LC system such that a sample is being loaded and desalted on one column whilst at the same time a different sample is being eluted from a second column and analysed by LC–MS/MS [10]. Nano Fraction Analysis Chip Technology (nanoFACT) was introduced as an alternative method for improving sample throughput and robustness [11]. Here the samples were automatically off-line separated by HPLC, fractionated and collected in the lumen of pipette tips. Each tip was subsequently applied to a miniaturised ESI chip emitter enabling mass spectrometric screening of the individual fractions.

In our research group efforts have been dedicated towards developing simple and robust peptide and protein sample preparation protocols suitable for MS analysis. Initially, peptide samples were desalted and concentrated using disposable miniaturised SPE reversed-phase column packed in the tip of GELoader tips [12,13]. In later developments the microcolumns were produced of small disks of chromatographic beads embedded in a Teflon meshwork [14,15].

In the present study we implemented an automated system that interfaces 'off-line' sample preparation using SPE disk columns to an 'on-line' LC–MS/MS system. We developed the novel SPE-LC design for fast, sensitive and robust peptide separation while operated at a high duty cycle (6–10 samples per hour). Initially, the samples are manually loaded onto SPE tips that are subsequently positioned in the 192 rack autosampler robotics. The peptides are recovered from the SPE precolumn and separated on the analytical column by isocratic elution in an automated workflow, allowing unattended analysis of 192 samples.

The chromatographic performance of the SPE-LC system was evaluated by analysing varying amounts of tryptic bovine serum albumin (BSA) and assessing the separation capacity in terms of peptide ion peak widths, column peak capacity and retention time reproducibility. Furthermore, it was established that SPE-LC–ESI-MS/MS provides sufficient resolving power and adequate peptide sequence information for the identification of minute amounts of protein by means of automated data acquisition, processing and annotation.

2. Experimental

2.1. In-solution protein digestion

The following proteins, alone or in a 12 component mixture were used for testing the SPE-LC–MS/MS setup: Transferrin (human) (gift from ACE Biosciences, Odense, DK), serum albumin (bovine), β-lactoglobulin (bovine), carbonic anhydrase (bovine), β-casein (bovine), α-casein (bovine), ovalbumin (chicken), ribonuclease B (bovine pancreas), alcohol dehydrogenase (bakers yeast), myoglobin (whale skeletal muscle), lysozyme (chicken) and α-amylase (Bacillus subtilis) (Sigma–Aldrich, St. Louis, MO). Each protein was dissolved in 0.1 M NH_4HCO_3 and 10 mM dithiothreitol (DTT) for reduction (1 h, 56 °C). Next, the solution was incubated with 50 mM iodoacetamide in the dark for 30 min at room temperature for alkylation. The proteins were digested using trypsin (Promega, Madison, WI; 1%, w/w; 37 °C over night) and the reaction was quenched by acetic acid (AcOH; 0.5%), upon which aliquots were stored at −20 °C.

2.2. In-gel protein digestion

In-gel digestion of protein was performed according to Shevchenko et al. [16]. Silver stained protein spots were excised from the gels, cut into small pieces and washed several times with 50 mM NH_4HCO_3 in 50% acetonitrile (MeCN). After complete destaining, the gel was washed in water and MeCN. The supernatant was removed and the gel pieces were rehydrated in 20 μL of 10 mM DTT and 50 mM NH_4HCO_3 in 1 h at 56 °C for reduction. The supernatant was removed and replaced with 20 μL 55 mM iodoacetamide and incubated in the dark for 30 min at room temperature for protein alkylation. The supernatant was removed and the gel pieces were shrunk in MeCN. After removal of the supernatant the gel pieces were rehydrated in a digestion buffer containing 50 mM NH_4HCO_3, and 5 ng/μL of trypsin (Promega, Madison, WI) and incubated at 37 °C over night.

2.3. SPE-LC separation

An aliquot of each protein digest was dissolved in 20 μL 0.5% AcOH in water, loaded onto disposable reversed-phase C18-p200 SPE tips (Proxeon Biosystems, Odense, Denmark) and de-salted with 30 μL 0.5% AcOH according to Rappsilber et al. [14].

Subsequently, the SPE tips were mounted in-line with a prototype LC system and the peptides separated by reversed-phase chromatography using isocratic elution. Flow was delivered by two nanoflow piston pumps operated with advanced flow control (AFC; Proxeon Biosystems, Odense, Denmark) and connected to a steel capillary (OD 1/32″, ID 125 μm; Upchurch Scientific, Oak Harbor, WA) that allowed a tight fit inside the SPE tip just above the adsorbing material. The tips were moved from their

rack position and held in a funnel shaped connector using a modified autosampler, thereby providing a flow path to the analytical column (ReproSil-Pur 120 AQ-C18, 3 μm; Dr. Maisch GmbH, AmmerbuchEntringen, Germany) packed in a 6 cm 75 μm ID fused silica emitter (Proxeon Biosystems, Odense, Denmark). A-solvent was 30% MeCN, 0.5% AcOH in water; B-solvent was 80% MeCN and 0.5% AcOH. A-solvent was initially passed through the columns at a flow rate of 600 nL/min for 2 min. Next, the flow rate was reduced to 30 nL/min for 5 min and in the end of the sequence, the analytical column was flushed with B-solvent at a flow rate of 1500 nL/min.

2.4. LC separation

Reversed phase LC was performed using the Proxeon EASY-nLC™ (Proxeon Biosystems, Odense, Denmark). The analytical column (ReproSil-Pur 120 AQ-C18, 3 μm; Dr. Maisch GmbH, AmmerbuchEntringen, Germany) was packed in a 8 cm 50 μm ID fused silica with a fritted plug. The flow rate was 200 nL/min and the peptides were separated by gradient elution from 0% to 40% B-solvent in 30 min. A-solvent was 0.5% AcOH in water; B-solvent was 80% MeCN and 0.5% AcOH.

2.5. ESI-MS/MS

Mass spectra were acquired using a Q-Star XL ESI-MS/MS system (MDS Sciex, Toronto, Canada) operated in positive ESI mode. A potential difference between the MS orifice and the emitter of 2300 V was applied through a liquid junction. Each MS spectrum was acquired for 1 s and the three most abundant doubly or triply charged precursor ions were automatically selected for MS/MS by information dependent acquisition (IDA) (1 MS/MS spectrum per second). The MS spectra were acquired in the *m*/*z* range of 350–1400.

2.6. MALDI sample preparation

Each protein aliquot was dissolved in 20 μL 0.5% AcOH in water, loaded onto disposable reversed-phase C18-GELoader SPE tips (Proxeon Biosystems, Odense, Denmark) and desalted with 30 μL 0.5% AcOH according to Rappsilber et al. [14]. The purified and concentrated peptides were eluted from the microcolumn in several droplets directly onto a MALDI probe using α-cyano-4-hydroxycinnamic acid (Sigma–Aldrich, St. Louis, MO) matrix solution (10 mg/mL in 50% MeCN/0.1% TFA).

2.7. MALDI-MS/MS

MALDI spectra were acquired on the 4700 Proteomics Analyzer MALDI-TOF/TOF system (Applied Biosystems, Foster City, CA) equipped with a 355 nm solid-state laser (diode pumped Nd:YAG laser) pulsing at a repetition rate of 200 Hz. In MS/MS mode, a collision energy of 1 kV was applied and nitrogen was used as collision gas. The spectra were acquired in the *m*/*z* range of 700–4500.

2.8. Protein database searching

ESI-MS/MS Mascot generic files were compiled from wiff files using the script Mascot.dll embedded in the Analyst QS 1.1 software (MDS Sciex, Toronto, Canada). MALDI-MS and MS/MS Mascot generic files were compiled from T2D files using in-house scripts embedded in Data Explorer (Applied Biosystems, Foster City, CA). Protein identification was performed using an in-house Mascot server (Matrix science Ltd, London, UK) applying the following settings; MS accuracy 50 ppm, MS/MS mass accuracy 0.2 Da, trypsin digestion with one allowed missed cleavage, fixed modification (C) carbamidomethyl, variable modification (M) oxidation. In-silico protein digestions and MS/MS sequence analysis were performed using GPMAW 7.0 (Lighthouse data, Odense, Denmark).

3. Results

3.1. Description of the SPE-LC platform

Most commercially available nanoflow LC systems are developed to achieve the highest degree of sample separation, whereas the main motivation driving our SPE-LC design is to increase sample throughput by decreasing a complete analysis cycle time to less than 10 min and at the same time maintaining a sufficient separation power and sensitivity to analyse samples of medium complexity by ESI-MS/MS.

The SPE sample loading, desalting and concentration step was performed off-line, and the SPE tips were positioned in the SPE-LC autosampler (192 positions) for automated analysis. The SPE-LC setup was mainly built from standard LC components; nanoflow pumps and valves assembled in a novel configuration (Fig. 1A). The increased analysis speed was primarily caused by two factors: (i) the decoupling of the sample loading and washing procedures on the SPE precolumns from the actual chromatographic separation process; and (ii) the use of isocratic elution instead of gradient elution. Using the modified autosampler, a SPE precolumn was moved from the 192-tip rack and connected in-line to a 6 cm analytical column, packed inside a pulled fused silica electrospray-emitter. By means of isocratic elution (solvent A) the peptides were recovered from the SPE precolumn and separated on the analytical column. Upon completion of the analysis, the analytical column was washed (solvent B) and the SPE tip was discarded.

First, the optimal solvent conditions for isocratic elution from the SPE C18 tips were investigated and optimised. The experiment was conducted manually by off-line loading and stepwise isocratic elution (using air pressure supplied by a syringe) of the C18 SPE tip followed by MALDI-MS investigation of the collected fractions. For each step, the MeCN ratio of the elution solvent was increased by 10% going from 0% to 100% in 11 stages. Five picomole tryptic BSA peptides were retained on the SPE column in the 0% MeCN step and eluted efficiently using a solvent composition of 0.5% AcOH in 30% MeCN (data not shown). Thus, 30% MeCN was used in all subsequent experiments for isocratic elution.

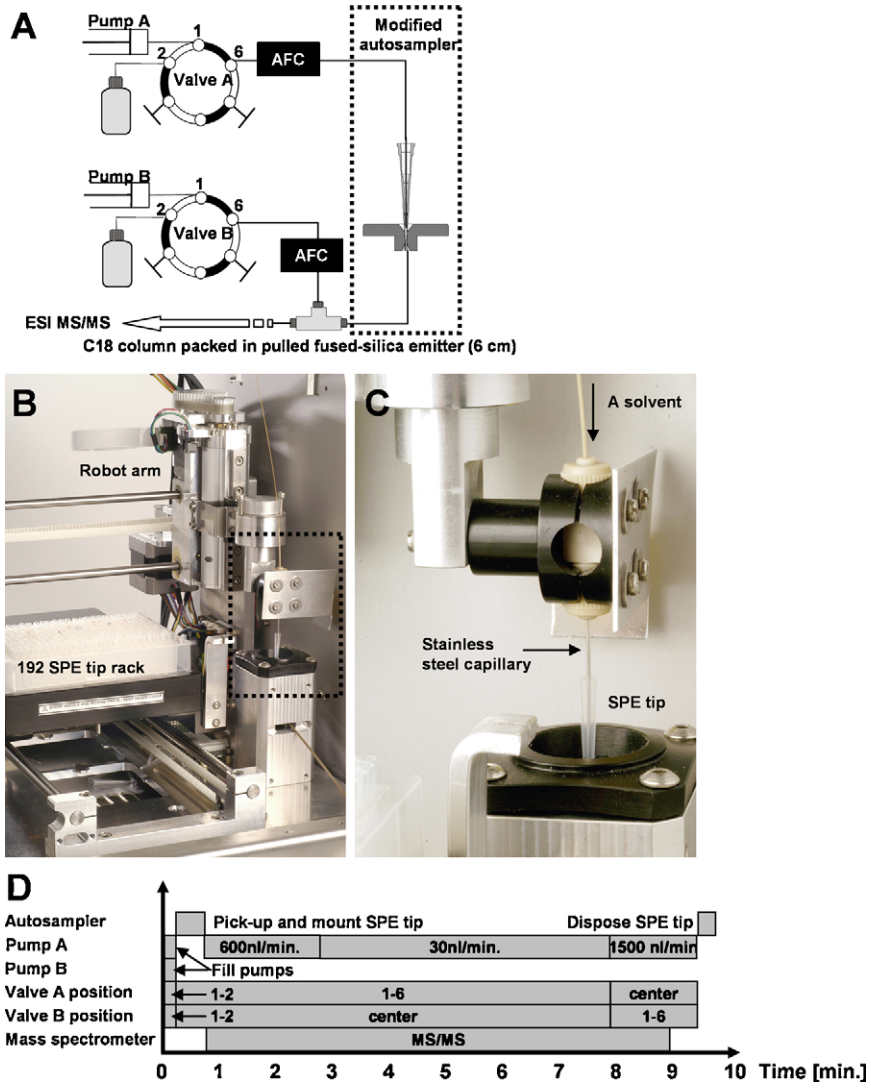

Fig. 1. The SPE-LC setup and operation. (A) Diagram of the SPE-LC system. Flow is delivered by two nanoflow piston pumps operated with advanced flow control (AFC). Pump A is connected to a stainless steel capillary mounted on a modified autosampler. The capillary allows a tight fit inside the SPE tip just above the adsorbing material, allowing a liquid connection from the SPE precolumn to the analytical column. Peptides are eluted from the SPE tip to the analytical column and separated using isocratic elution (30% MeCN) delivered by pump A. (B) Picture of the modified autosampler. The robot arm picks up an SPE tip from the rack using the stainless steel capillary, moves the tip to the peek adapter and seals the connection by applying downward mechanical pressure. After the analysis is concluded, the robot arm disposes of the SPE tip. (C) Close up picture of the interface between the SPE tip and the LC system. (D) Operation chart for the key-components of the SPE-LC system. The different components are represented on the *y*-axis and the time in minute scale is depicted along the *x*-axis. Every analysis begins with pump refilling and pressure calibration. Next, a SPE precolumn is picked up and mounted in the adaptor by the autosampler robot arm and a contact closure signal is sent to the mass spectrometer. Pump A delivers flow at 600 nL/min. After 2 min the flow rate is lowered to 30 nL/min for 5 min. Next, Pump B delivers 1500 nL/min for 3 min and finally the SPE precolumn is disposed off. Valves in position 1–2 connect the piston pumps to the solvent containers, in position 1–6, the pumps connect to the SPE-LC system and in center position the pumps are blocked.

We note that many detergent and polymer contaminants were still effectively retained on the disposable C18 SPE precolumn at 30% MeCN, and thus the adverse effects of these contaminants are greatly reduced when performing separations with organic phase at this ratio. This is important as detergents and polymers reduce the chromatographic performance, the analytical column lifetime and they generate multiply charged ions in ESI, which are often selected for MS/MS analysis but yield no useful information.

Next, the SPE-LC flow rate was optimised to achieve relative high throughput while maintaining high sensitivity LC-MS/MS. A high flow rate ensures rapid sample elution, whereas a low flow rate provides more time for MS/MS data acquisition which in turn leads to higher quality mass spectra.

In the first experiments, aimed at testing the performance of the setup, a constant flow rate was applied to the LC system. A sample of 100 fmol of tryptic BSA peptides was loaded onto a SPE tip, washed, and the tip was transferred to the SPE-LC autosampler and automatically connected in-line to the analytical column. A solvent composed of 30% MeCN and 0.5% acetic acid was passed through the SPE tip and the analytical column at a flow rate of 300 nL/min and analysed by ESI-MS/MS. In order to evaluate the chromatographic separation in terms of peak widths and the column capacity, the base peak chromatogram

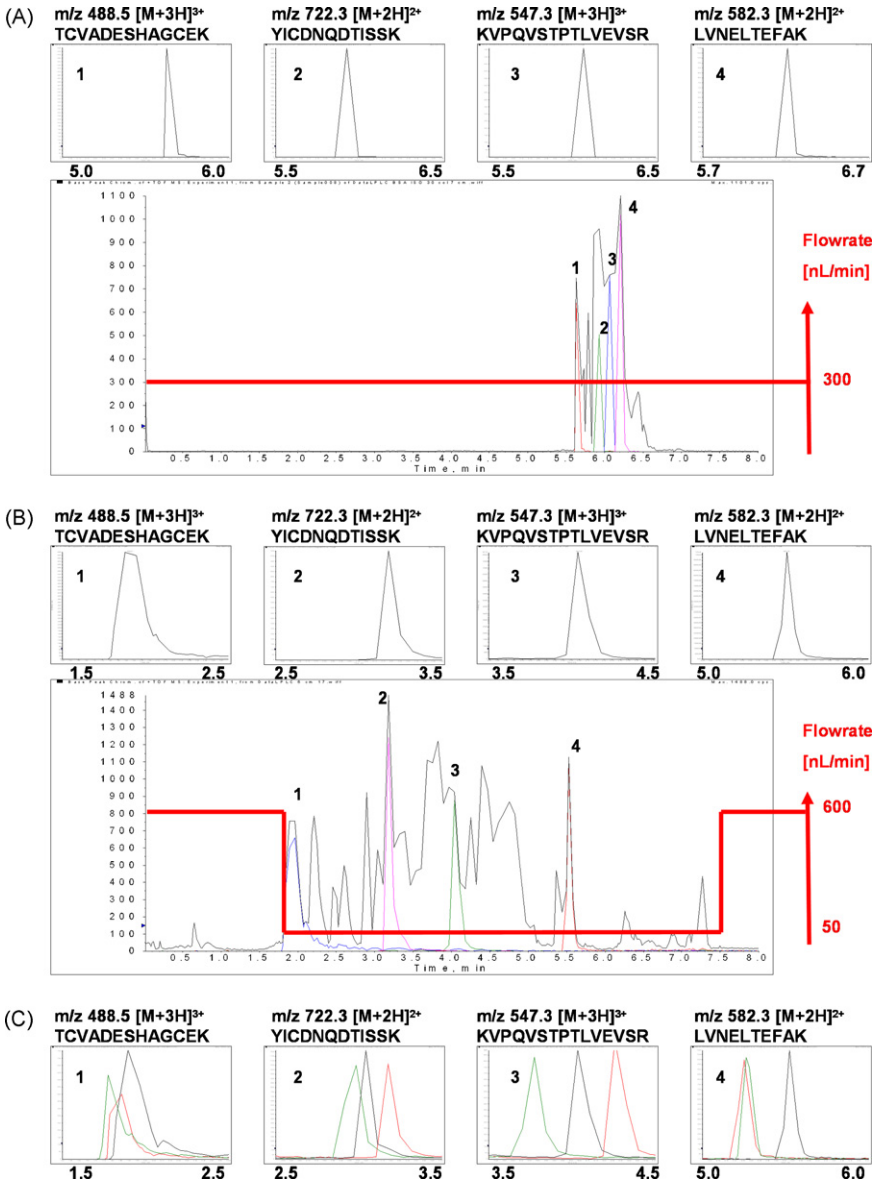

Fig. 2. SPE-LC–ESI-MS chromatographic performance for the analysis of 100 fmol BSA. (A) Isocratic constant flow (300 nL/min) SPE-LC–ESI-MS/MS analysis of 100 fmol tryptic BSA: base peak chromatogram with overlaid extracted ion chromatograms of four selected ions. Inserts show 1 min zoom sections of the four extracted ions chromatograms. (B) Isocratic dynamic flow SPE-LC–ESI-MS/MS analysis of 100 fmol tryptic BSA: flow profile and base peak chromatogram with overlaid extracted ion chromatograms of four selected ions. Inserts show 1 min zoom sections of the four extracted ions chromatograms. (C) three consecutive SPE-LC–ESI-MS/MS analyses of 100 fmol BSA using the dynamic flow profile depicted in Fig. 2B. The extracted ion chromatograms of four selected ions are overlaid in 1 min zoom sections.

(BPC) and extracted ion chromatograms (XIC) of selected peptide ions were investigated (Fig. 2A). *m/z* 488.5 ($[M+3H]^{3+}$, 76-TCVADESHAGCEK-88) was the first identifiable peptide ion and it eluted after 5.6 min whereas the last eluting peptide at *m/z* 507.8 ($[M+2H]^{2+}$, 549-QTALVELLK-557) was detected after 6.3 min, which means that all peptides eluted within 0.7 min corresponding to 42 s (t). Assuming that the data is normally distributed, the peak mean ± 2 S.D. (σ) account for 95.4% of the data points. The peak width at the 4σ level can be measured at 13.4% of the maximum peak intensity. The average peak width of the 10 most abundant peptide ions was measured to be 8 s at the 4σ level, however this number may be slightly exaggerated as the Q-TOF mass spectrometer was operated in information dependent acquisition (IDA) mode with an MS and MS/MS operation cycle of 4 s and therefore only collecting a limited number of MS data points across the eluting peptide peaks. The theoretical column peak capacity (n_c) is a measure of the number of peaks that can be separated to the 4σ level within the peptide elution window. $n_c = t/4\sigma$ [17], which in our case means $n_c = (42\text{ s}/8\text{ s}) = 6$. Thus, six peaks can be separated to the 4σ level during the 0.7 min of peptide elution. In summary, using constant flow rate (300 nL/min) isocratic elution SPE-LC all peptides eluted within 0.7 min having individual peak width of 8 s (at the 4σ level).

The current mass spectrometer settings allow acquisition of three MS/MS spectra in 4 s, thus 32 peptides can be sequenced

during the 42 s of SPE-LC–MS/MS analysis. According to statistics on the UniProtKB/Swiss-Prot protein knowledgebase release 52.0, the average protein is comprised of 365 amino acids of which 11.34% are either arginine or lysine. Thus tryptic proteolysis will yield on average 40 peptides per protein (365×0.1134) assuming complete cleavage. Taken together the current SPE-LC–ESI-MS/MS method will allow comprehensive analysis of 1–2 proteins if all theoretical peptides are generated and ionised in ESI. In order to analyse samples of medium complexity (3–12 proteins), comprising on the order of 100–500 peptides, the initial SPE-LC method was further enhanced.

Additional optimisation was achieved by increasing the duration of the peptide elution window and reducing the total analysis time by using a variable flow rate over the course of the SPE-LC analysis. The flow rate was low (30 nL/min) during peptide elution and high (600 nL/min) at other time points of the analysis (Fig. 2B). Nanoelectrospray (flow rate ~25 nL/min, emitter orifice ID ~ 1–3 μm) is known to provide high ionisation efficiency [18,19]. The spraying conditions of SPE-LC (flow rate 30 nL/min, emitter orifice ID ~ 6 μm) are similar to those of nanoelectrospray, and thus the ionisation efficiency is expected to be comparable as well.

The optimised variable flow isocratic elution SPE-LC method was tested by analysing 100 fmol of tryptic BSA peptides. Separation efficiency was evaluated in terms of peptide peak width and column capacity. The flow profile and extracted ion chromatograms from the analysis are depicted in Fig. 2B. m/z 488.5 ($[M+3H]^{3+}$, 76-TCVADESHAGCEK-88) eluted after approximately 2 min, whereas the last eluting peptide m/z 507.8 ($[M+2H]^{2+}$, 549-QTALVELLK-557) was detected after 6.3 min. The entire peptide content eluted within 4.3 min ($t = 258$ s) and the average peak width at the 4σ intensity level of the 14 most abundant ions was measured to be 12 s, yielding a peak capacity of, $n_c = (258\text{ s}/12\text{ s}) = 22$. Using these settings ~190 peptides can be sequenced during a 258 s SPE-LC–MS/MS run, which will easily enable analysis of medium complex tryptic protein mixtures (3–12 proteins corresponding to 100–500 peptides). It should be noted that newer mass spectrometers are considerable faster than the instrument used in the present experiments and coupling the SPE-LC system to a modern MS could enable the acquisition of more than 1000 MS/MS spectra in 5 min.

The narrow peak width and the low flow rate of 30 nL/min, means that a 100 fmol sample (vol. = 1 μL) elutes at an average peptide ion concentration of $0.1\text{ pmol}/(0.2\text{ min} \times 0.03\ \mu\text{L/min}) \sim 16\ \text{pmol}/\mu\text{L}$, which corresponds to a concentration factor of $(16\ \text{pmol}/\mu\text{L})/(0.1\ \text{pmol}/\mu\text{L}) = 160$. Thus, using the variable flow profile method, the high throughput was maintained while the peptides eluted over a longer time span thereby facilitating more MS/MS spectra acquisitions, and the separation was improved in terms of column capacity. Hence, variable flow isocratic elution was used for the remaining experiments described herein.

Having established a protocol for running SPE-LC experiments, the chromatographic run-to-run reproducibility of the system was examined. For this investigation, three identical samples of 100 fmol BSA were analysed in succession by SPE-LC–ESI-MS. The extracted ion chromatograms of four selected peptide ions were compared (Fig. 2C). The average retention time standard deviation (SD) of the four peaks was 0.12 s and the average relative standard deviation (RSD) was 3.4%. This performance might be limiting for a range of applications, e.g., quantification experiments based on run-to-run comparisons of extracted ion intensities. For many other applications a retention time RSD of around 3% is quite adequate. For example, quantification based upon stable isotope labelling of peptides is not negatively affected because the isotope labelled peptides co-elute regardless of any jitter. One of the major factors affecting the SPE-LC retention time reproducibility is geometry variations of the disposable SPE precolumns that are currently not manufactured in a sufficiently reproducible manner. New procedures for a more consistent production are being investigated with the aim to alleviate this problem.

Table 1
Chromatographic parameters for SPE-LC and conventional capillary LC separation of 100 fmol tryptic BSA digest

	SPE-LC	Capillary LC
Flow rate (nl/min)	30	200
Peptide elution (s)	258	1200
Peak width at 13.4% (s)	12	20
Column capacity	22	60
Average peak concentration (pmol/μl)	16	1.5
Concentration factor	160	15
Retention time RSD (%)	3.4	~0.2
Throughput (samples/h)	6–10	1–2

The SPE-LC was operated with the isocratic dynamic flow elution as described in Fig. 2B, whereas the capillary LC was operated with gradient elution from 0% to 40% solvent B in 30 min.

The separation performance of the SPE-LC system was demonstrated by analysis of simple samples. The key separation parameters of SPE-LC and conventional capillary LC were compared. SPE-LC–MS/MS and capillary LC-MS/MS data were obtained from identical peptide samples (100 fmol BSA, tryptic digestion) and summarised in Table 1. According to these results, conventional LC can separate approximately three times more components (to the 13.4% intensity level) and provides better reproducibility than the SPE-LC setup. On the other hand, SPE-LC gives a 10 times higher average concentration of the individual species over the course of elution and it is working at a significant higher sample throughput. Thus, the SPE-LC system has exactly the features and performance needed for routine analysis of low- to semi-complex peptide samples. The potential of SPE-LC–ESI-MS/MS as a proteomics tool for protein identification and characterisation is further demonstrated below.

3.2. Proteomics applications of SPE-LC–ESI-MS/MS

The sensitivity of the SPE-LC–ESI-MS/MS setup for protein characterisation was evaluated by triplicate analysis of varying amounts of tryptic BSA samples (100, 50 and 10 fmol). In total, nine SPE precolumns were mounted in the autosampler and automatically analysed over the duration of 94 min (average 10.4 min per sample). The Q-TOF-MS/MS spectra were pro-

Table 2
SPE-LC–ESI-MS/MS protein identification sensitivity test

BSA sample	10 fmol	10 fmol	10 fmol	Average 10 fmol	50 fmol	50 fmol	50 fmol	Average 50 fmol	100 fmol	100 fmol	100 fmol	Average 100 fmol
Mascot score	1011	1024	1068	1034	1645	1638	1653	1645	1756	1641	1541	1646
Sequence coverage	38	39	37	38 (31)	51	48	50	50 (44)	52	48	51	50 (46)
No. of peptides	23	24	23	23	35	35	37	36	38	38	38	38

Triplicate SPE-LC–ESI-MS/MS analysis of 10, 50 and 100 fmol tryptic BSA followed by automated data processing and Mascot database searching. For each record, the Mascot score, the BSA protein sequence coverage and the number of peptides used for identification is listed. For every triplicate experiment the overlapping BSA sequence coverage is listed in parenthesis.

cessed (Analyst QS 1.1 script) and annotated using the Mascot database search engine. In every instance BSA was confidently identified as summarised in Table 2.

At a sample load of 10 fmol, BSA was identified with an average Mascot score of 1034 and a protein sequence coverage of 38%. Furthermore, the reproducibility of the MS/MS sequence coverage was high. In total, 31% of the BSA sequence was covered in all three experiments. Four MS/MS spectra of selected peptide ions from a 10 fmol BSA sample analysis are presented in Fig. 3. In every instance, peptide sequence ladders consist-

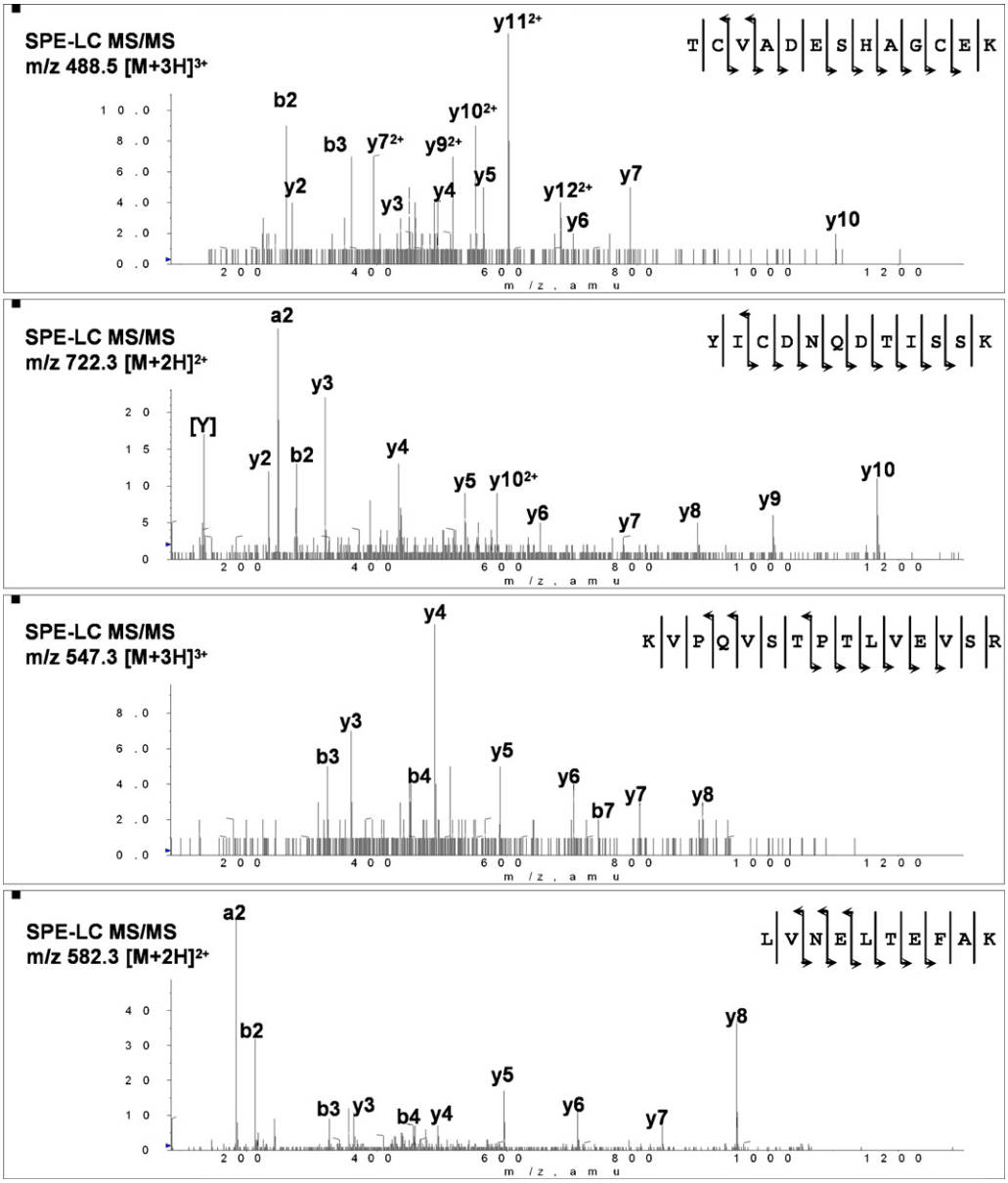

Fig. 3. Four MS/MS spectra from the SPE-LC–ESI-MS/MS analysis of 10 fmol of tryptic BSA. Immonium ions are marked by capital letters, whereas sequence specific b and y ions are marked in lower case.

Table 3
Protein Mascot scores resulting from the analysis of a 12-component mixture by conventional capillary LC–ESI-MS/MS and SPE-LC–ESI-MS/MS

Protein	In-silico tryptic peptides[a]		LC–ESI-MS/MS (gradient)		SPE-LC–ESI-MS/MS (isocratic)		SPE-LC–ESI-MS/MS (two runs combined)	
	All	700–5000 Da	Mascot	Peptides[a]	Mascot	Peptides[a]	Mascot	Peptides[a]
Serum albumin	82	45	1571	31	1016	26	1249	28
Transferrin	80	43	1594	31	864	19	1082	22
β-Lactoglobulin	19	10	341	6	298	6	356	7
Alcohol dehydrogenase	31	18	640	11	313	7	360	9
Lysozyme	19	10	352	6	286	5	322	5
Carbonic anhydrase	25	15	364	7	172	4	294	7
β-Casein	17	8	103	2	103	2	103	2
Ovalbumin	34	19	322	5	–		125	3
Myoglobin	22	10	188	5	108	3	169	6
α-S1-casein	21	13	78	2	106	3	106	3
α-S2-casein	31	16	265	5	–		61	3
Trypsin	15	13	87	2	156	2	156	2
Ovomocoid precursor	17	11	118	3	77	2	77	2
Ubiquitin	12	6	143	4	73	2	106	3
α-Amylase	57	37	–		–		–	
Ribonuclease B	16	7	–		–		–	
Total	498	261	6048	120	3429	81	4570	102
Average per protein	31	16	432	7.8	286	6.8	326	7.3
Average per peptide			50		42		45	

The data were acquired, processed and database searched in an automated workflow. The theoretical number of peptides per protein was determined by in-silico trypsin digestion, and the results presented in the first column. The results of the capillary LC–ESI-MS/MS and the SPE-LC–ESI-MS/MS analysis are presented in the second and third columns, whereas the results from two combined SPE-LC–ESI-MS/MS datasets are shown in the fourth column.

[a] The in-silico computation is based upon complete tryptic digestion, whereas the experimentally identified unique peptides were allowed one missed cleavage.

ing of at least six consecutive y-ion fragments were identified. Using either 50 or 100 fmol BSA, an average MS/MS sequence coverage of approximately 50% was observed.

An in-silico tryptic BSA digest, revealed that most of the undetected peptides were below a size of 700 Da or above 2000 Da. This bias was probably caused by the SPE-LC–ESI-MS/MS setup; peptides below 700 Da were not detected (as multiply charged species) by the mass spectrometer, whereas the larger and hydrophobic peptides were likely retained on the C18 SPE precolumn due to the isocratic elution solvent composition of 30% MeCN. This might also explain the non-linear gain in sequence coverage going from 50 to 100 fmol. It does not seem to be caused by mass spectrometer detector saturation or diminished chromatographic performance due to sample overload. However, almost every detectable peptide is observed from the 50 fmol sample, and hence doubling the sample load to 100 fmol does not improve the result significantly. The coinciding BSA sequence coverage of the triplicate experiments using 50 and 100 fmol BSA tryptic digest is 42% and thus the MS/MS peptide ion selection was nearly identical independent of the sample load.

In conclusion, as little as 10 fmol of tryptic BSA could be confidently identified by SPE-LC–MS/MS and furthermore, the reproducibility in terms of MS/MS precursor ion selection and amino acid sequence coverage obtained in triplicate experiments was high.

These results were based upon analysis of simple protein samples with only one component, but if the setup should be able to bridge the gap between MALDI-MS/MS for simple peptide mixtures and LC–MS/MS for complex samples, the SPE-LC–MS/MS system must be able to analyse medium complexity samples (3–12 protein components), e.g., from SDS–PAGE gel bands.

To further investigate the potential of the SPE-LC–MS/MS, a 12 component protein sample was analysed by direct MALDI-MS/MS (with prior off-line SPE desalting step), SPE-LC–MS/MS and capillary LC–MS/MS. In-silico tryptic digest of the 12 component protein mixture computed 498 tryptic peptides (assuming complete cleavage) which corresponded to 31 peptides per protein, on average. This was slightly less than predicted based upon the UniProtKB/Swiss-Prot database statistics. The proteins from the mixture were, however, smaller than the database average, and thus the number of peptides per protein was lower. Furthermore, only 261 of the 498 peptides were within the mass spectrometer detection range (700–5000 Da). The results obtained by SPE-LC–MS/MS and capillary LC–MS/MS of this tryptic digest analyses were processed and annotated using identical software parameters and the Mascot database search results are presented in Table 3.

Capillary LC–ESI-MS/MS analysis of the protein mixture resulted in the identification of 10 of the 12 known proteins. α-Amylase and Ribonuclease B were not detected. Ribonuclease B is a poor substrate for trypsin, in part due to an N-linked glycan at position N-60 that creates steric hindrance, which can explain why it was not detected. Four additional proteins were identified from the protein mixture; α-casein was a mixture of S1 and S2 α-casein, ovomocoid precursor ('contamination' of ovalbumin), ubiquitin and trypsin (autolysis). Thus, capillary LC–MS/MS identified a total of 14 proteins based upon 120 unique peptide

Table 4
Direct MALDI-MS/MS and SPE-LC–ESI-MS/MS analysis of porcine fetal neural stem cell proteins separated by 2DE

Sample	Protein ac.	Protein name	Method	Mascot	Peptide MS/MS	MS/MS seq. cov.
4019	Q5E946	DJ-1 protein (*Bos taurus*)	MALDI MS and MS/MS	65	0	0
			SPE-LC ESI MS/MS	192	4	24
5326	Q15084	Protein disulfide-isomerase A6 precursor (*Homo sapiens*)	MALDI MS and MS/MS	68	1	4
			SPE-LC ESI MS/MS	352	5	9
7407	P41367	Medium-chain specific acyl-CoA dehydrogenase (*Sus scrofa*)	MALDI MS and MS/MS	198	4	12
			SPE-LC ESI MS/MS	362	9	19
7725	GI:57209130	Paraspeckle component 1 (*Homo sapiens*)	MALDI MS and MS/MS	57	1	2
			SPE-LC ESI MS/MS	76	2	6
7741	Q12828	Far upstream element-binding protein 1 (*Homo sapiens*)	MALDI MS and MS/MS			
			SPE-LC ESI MS/MS	169	4	5

For each protein hit the Mascot score, protein sequence coverage and number of peptides used for identification are listed for both the MALDI-MS/MS and the SPE-LC–ESI-MS/MS analysis.

annotations from 596 MS/MS spectra with an average Mascot score of 50. Fourty-six percent of the theoretically detectable peptides were thus identified using capillary LC–MS/MS.

The SPE-LC–ESI-MS/MS data contained sufficient information to identify 9 of the 12 expected proteins plus three additional proteins. Eighty-one unique peptides, corresponding to 31% of the theoretically detectable peptides, were identified with an average Mascot score of 44. Thus, the average peptide Mascot score is similar for the capillary LC–MS/MS and SPE-LC–MS/MS analysis, but two proteins less were identified and the average protein sequence coverage was lower. However, the limiting factor for the SPE-LC–MS/MS analysis was the MS/MS acquisition rate of the Q-TOF mass spectrometer. During the approximately 5 min of peptide elution, the mass spectrometer was acquiring the maximum number of MS/MS spectra possible, but even so, many potential precursor ions were missed. To compensate for this factor, the data from two consecutive SPE-LC–MS/MS analysis were combined. In this way, 10 of the 12 expected plus four additional proteins were identified, so the total number of identifications equalled that achieved by LC–ESI-MS/MS. Still, α-amylase and ribonuclease B remained undetected as in the LC–ESI-MS/MS analysis. In total, 41% of the theoretically detectable peptides were identified. Furthermore, the two SPE-LC–ESI-MS/MS dataset were largely overlapping, which reflects the good reproducibility. Nevertheless, potentially an even better result could have been achieved by using an ion exclusion list. In addition, the result can be improved by using a mass spectrometer with faster data acquisition. New generations of Q-TOFs and other hybrid tandem mass spectrometers produce very accurate MS mass measurements (mass error below 10 ppm) and can acquire MS/MS spectra at a higher rate (e.g., 4–8 MS/MS spectra per second).

Direct MALDI-MS and MS/MS analysis of the desalted peptide sample only identified 4 of the 12 expected proteins. The 50 most intense precursor ions were selected for MS/MS acquisition resulting in 15 peptide identifications with Mascot scores above 20. After the ∼20 first MS/MS spectra were acquired the quality deteriorated. The amount of sample left on the MALDI target was simply too low and prohibited a more detailed analysis. Furthermore, TOF–TOF type instruments are not able to select precursor ions at unit resolution and, due to the sample complexity, several precursor ions were simultaneously collisionally activated per MS/MS experiment, which resulted in very complex MS/MS spectra. Thus, it was only possible to identify five proteins using MALDI MS and MS/MS.

In summary, the SPE-LC–MS/MS setup enabled analysis and identification of up to 12 proteins from one sample. This performance is better than what was achieved by direct MALDI-MS and MS/MS analysis, but lower than by conventional capillary LC–MS/MS. The MALDI-MS/MS and two SPE-LC–MS/MS analysis both required 20 min duration, whereas 1 h was spent on the LC–MS/MS analysis.

Finally we investigated the performance of the SPE-LC–MS/MS platform for the analysis of biologically relevant samples using a 2DE proteomics approach. Comparative proteomic analysis of steady state and differentiating porcine fetal neural stem cells was investigated using an in vitro model system [20]. Proteins from fetal stem cells and their differentiated progenies were separated by 2DE and quantified based upon spot intensity. Differentially expressed proteins were detected and the corresponding protein spots cut out from the gel. Proteins were digested using trypsin, the peptides extracted and each resulting sample was split in two. One aliquot was analysed by MALDI-MS and MS/MS (with prior desalting by SPE) and the other aliquot was analysed by SPE-LC–ESI-MS/MS. The results are presented in Table 4.

MALDI-MS and MS/MS identified four proteins whereas SPE-LC–MS/MS identified five different proteins. In every case, the SPE-LC–ESI-MS/MS method provided more confident data than MALDI-MS and MS/MS for the subsequent protein identification in terms of Mascot score and MS/MS sequence coverage.

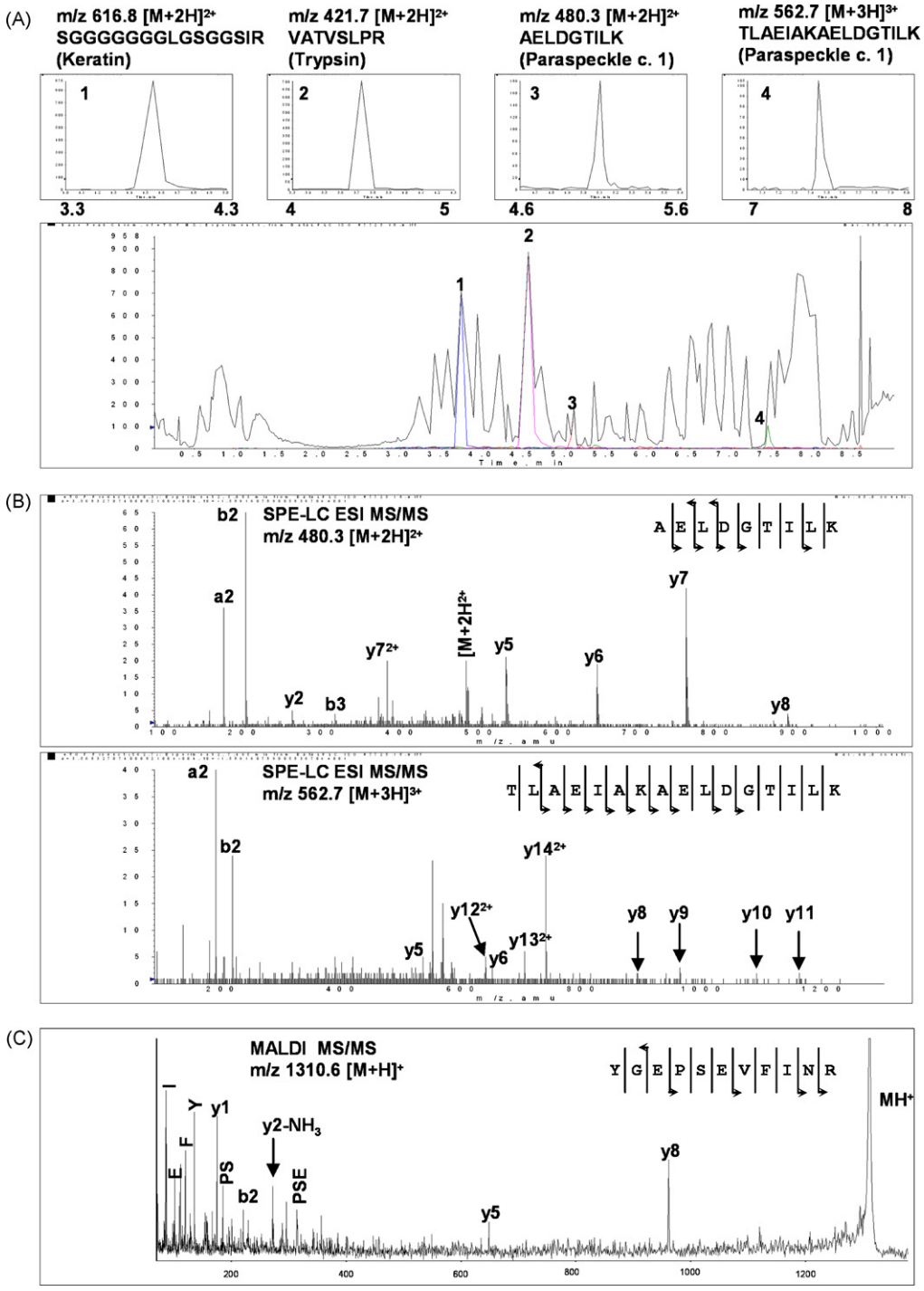

Fig. 4. SPE-LC–ESI-MS/MS analysis of 2DE protein spot #7725 from porcine fetal neural stem cell. The 2DE gel spot was excised, protein in-gel digested by trypsin, peptides extracted and the sample split in two. One aliquot was analysed by MALDI MS and MS/MS and the other aliquot was analysed by SPE-LC–ESI-MS/MS. (A) Base peak chromatogram overlaid with the extracted ion chromatograms of selected peptide ions from the SPE-LC–ESI-MS/MS analysis. Peaks marked 1 and 2 are peptide ions from the known contaminants keratin and trypsin, respectively, whereas the less intense peaks marked 3 and 4 are paraspeckle component 1 peptide ions. Inserts show 1 min zoom sections of the four extracted chromatograms. (B) ESI-MS/MS spectra of *m/z* 480.3 ($[M+2H]^{2+}$) and *m/z* 562.7 ($[M+3H]^{3+}$) used for paraspeckle component 1 identification. (C) MALDI MS/MS spectrum of *m/z* 1310.6 ($[M+H]^{+}$) from paraspeckle component 1. Immonium and internal fragment ions are marked by capital letters, whereas sequence specific b- and y-fragment ions are marked in lower case.

Spot #7725 was the faintest of the selected protein spots and the top four protein hits from the Mascot search were keratins and trypsin, whereas paraspeckle component 1 came up as entry number 5. The base peak chromatogram and the extracted ion chromatograms of selected ions from the SPE-LC–ESI-MS/MS analysis are presented in Fig. 4A. Peaks 1 and 2 correspond to the extracted ion MS traces of a trypsin and a keratin peptide ion, respectively. The two peptide ions used for paraspeckle component 1 identification are marked peaks 3 and 4 in the base peak chromatogram. Based upon peak area integration of the four

peaks presented, it is estimated that paraspeckle component 1 was present in a ratio of 1:20:20 in the favour of trypsin and keratin.

Manual interpretation of the two paraspeckle component 1 MS/MS spectra, corresponding to two partly overlapping peptides (amino acids 128–143) (Fig. 4B), confirmed the initial Mascot ID. In both spectra the b- and y-fragment ion series covered more than 50% of the peptide sequence. Although, the intensities of some of the backbone fragment ions of the *m/z* 562.7 MS/MS spectrum were rather weak and some ions were doubly charged, their validity could be verified due to the high resolution (5000–8000 FWHM) and mass accuracy (RMS error < 120 ppm) of the Q-TOF instrument in MS/MS mode.

In addition, the SPE-LC data was complemented by a MALDI TOF/TOF tandem mass spectrum of *m/z* 1310.6 (Fig. 4C), which corresponded to a peptide ion from a different region of the paraspeckle component 1 sequence. The spectrum provided abundant information in the low mass region, corresponding to the expected immonium and internal fragment ions, whereas only a few backbone cleavage y-ions were present. However, the y5 and y8 fragment ions were theoretically assumed to be intense as they were formed by amide bond cleavage c-terminal to glutamic acid and between glutamic acid and proline, respectively.

In conclusion, paraspeckle component 1, along with four other proteins, separated by 2DE were unambiguously identified by SPE-LC–ESI-MS/MS. MALDI-MS and MS/MS supported these findings, but in every instance SPE-LC–ESI-MS/MS gave a more confident result in terms of Mascot score and MS/MS sequence coverage. This demonstrates that SPE-LC–ESI-MS/MS provides a rapid and sensitive platform for analysis of protein samples separated by 2DE, and that the system is a viable alternative to direct MALDI-MS/MS analysis.

4. Conclusion

The SPE-LC interface provides chromatographic characteristics such as; throughput of six samples per hour, peak capacity of 21, concentration factor 160 and an average run-to-run reproducibility of 3.4% retention time RSD. In combination with an ESI tandem mass spectrometer it can be used for confident identification of even minute amounts of protein from samples of medium complexity. A number of 2DE separated proteins were analysed by direct MALDI-MS/MS and SPE-LC–ESI-MS/MS, where the latter provided by far the most confident identifications.

The system is compatible with other orthogonal preseparation techniques in order to analyse complex samples, e.g., SDS–PAGE, strong cation exchange (SCX) and isoelectric focusing (IEF). In addition, particular classes of post translationally modified peptides like phosphopeptides, can be studied using, e.g., antibody based immunoprecipitation [21], Fe-IMAC [22–24] or titanium dioxide [25,26] followed by SPE-LC–MS/MS. In this way, SPE-LC–MS/MS is merely a rapid final separation and detection step that can be interfaced with the chromatographic method of choice for analysis of a broad range of protein samples.

Whereas the new SPE-LC technology already seems very powerful for proteomics analysis, future work will aim to improve the technology even further, where, based on current data, we hope to achieve a throughput of up to 12 samples per hour and a retention time reproducibility of around 1%.

Acknowledgements

We wish to thank Jens S. Andersen and Jakob Bunkenborg for access to and training on the Q-Star XL mass spectrometer. Furthermore, we thank Andrea Lorentzen, Helena Skalnikova, Petr Halada and Dr. Hana Kovarova for a successful collaboration on the stem cell project. Ole B. Hørning is supported by an industrial Ph.D. fellowship from Proxeon Biosystems A/S and the Danish Ministry of Science, Technology and Innovation. Ole N. Jensen is a Lundbeck Foundation Research Professor.

References

[1] R. Aebersold, M. Mann, Nature 422 (6928) (2003) 198.
[2] O.N. Jensen, Nat. Rev. Mol. Cell Biol. 7 (6) (2006) 391.
[3] A.J. Link, J. Eng, D.M. Schieltz, E. Carmack, G.J. Mize, D.R. Morris, B.M. Garvik, J.R. Yates, Nat. Biotechnol. 17 (7) (1999) 676.
[4] D.A. Wolters, M.P. Washburn, J.R. Yates, Anal. Chem. 73 (23) (2001) 5683.
[5] P.A. Haynes, N. Fripp, R. Aebersold, Electrophoresis 19 (6) (1998) 939.
[6] A.J. Link, L.G. Hays, E.B. Carmack, J.R. Yates, Electrophoresis 18 (8) (1997) 1314.
[7] S. Xie, R.W. Allington, F. Svec, J.M.J. Frechet, J. Chromatogr. A 865 (1–2) (1999) 169.
[8] Y. Shen, R.D. Smith, K.K. Unger, D. Kumar, D. Lubda, Anal. Chem. 77 (20) (2005) 6692.
[9] Y. Chen, S.C. Kim, Y. Zhao, Anal. Chem. 77 (24) (2005) 8179.
[10] E. Witters, K. Laukens, P. Deckers, W.V. Dongen, E. Esmans, H.V. Onckelen, Rapid Commun. Mass Spectrom. 17 (19) (2003) 2188.
[11] T.N. Corso, C.K. VanPelt, J. Li, C. Ptak, X. Huang, Anal. Chem. 78 (7) (2006) 2209.
[12] M. Kussmann, E. Nordhoff, H. Rahbek-Nielsen, S. Haebel, M.R. Larsen, L. Jakobsen, J. Gobom, E. Mirgorodskaya, A.K. Kristensen, L. Palm, P. Roepstorff, J. Mass Spectrom. 32 (6) (1997) 593.
[13] J. Gobom, E. Nordhoff, E. Mirgorodskaya, R. Ekman, P. Roepstorff, J. Mass Spectrom. 34 (2) (1999) 105.
[14] J. Rappsilber, Y. Ishihama, M. Mann, Anal. Chem. 75 (3) (2003) 663.
[15] A.K. Callesen, S. Mohammed, J. Bunkenborg, T.A. Kruse, S. Cold, O. Mogensen, R.D. Christensen, W. Vach, P.E. Jorgensen, O.N. Jensen, Rapid Commun. Mass Spectrom. 19 (12) (2005) 1578.
[16] A. Shevchenko, M. Wilm, O. Vorm, M. Mann, Anal. Chem. 68 (1996) 850.
[17] J.C. Giddings, John Wiley and Sons, INC., 1991; p. 105.
[18] M. Wilm, M. Mann, Int. J. Mass Spectrom. Ion Process. 136 (1994) 167.
[19] M. Wilm, M. Mann, Anal. Chem. 68 (1996) 1.
[20] H. Skalnikova, P. Halada, P. Vodicka, J. Motlik, P. Rehulka, O. Hørning, J. Chmelik, O.N. Jensen, H. Kovarova, Proteomics 7 (11) (2007) 1825.
[21] M. Gronborg, T.Z. Kristiansen, A. Stensballe, J.S. Andersen, O. Ohara, M. Mann, O.N. Jensen, A. Pandey, Mol. Cell. Proteomics 1 (7) (2002) 517.
[22] L. Andersson, J. Porath, Anal. Biochem. 154 (1986) 250.
[23] S.B. Ficarro, M.L. McCleland, P.T. Stukenberg, D.J. Burke, M.M. Ross, J. Shabanowitz, D.F. Hunt, F.M. White, Nat. Biotechnol. 20 (3) (2002) 301.
[24] A. Stensballe, S. Andersen, O.N. Jensen, Proteomics 1 (2001) 207.
[25] M.W.H. Pinkse, P.M. Uitto, M.J. Hilhorst, B. Ooms, A.J.R. Heck, Anal. Chem. 76 (14) (2004) 3935.
[26] M.R. Larsen, T.E. Thingholm, O.N. Jensen, P. Roepstorff, T.J.D. Jorgensen, Mol. Cell. Proteomics 4 (7) (2005) 873.