ELSEVIER

Journal of Chromatography A, 950 (2002) 99–111

JOURNAL OF CHROMATOGRAPHY A

www.elsevier.com/locate/chroma

Liquid chromatography–mass spectrometry and capillary electrophoresis combined approach for separation and characterization of multicomponent peptide mixtures Application to crude products of leuprolide synthesis ☆

V. Sanz-Nebot*, F. Benavente, J. Barbosa

Department of Analytical Chemistry, University of Barcelona, Av. Diagonal 647, 08028 Barcelona, Spain

Received 12 June 2001; received in revised form 31 December 2001; accepted 31 December 2001

Abstract

A sequential combination of reversed-phase liquid chromatography–mass spectrometry (LC–MS) and capillary electrophoresis (CE) has been explored in order to perform separation and characterization of a multicomponent peptide mixture from the synthesis of leuprolide. The mixture was first analyzed and fractionated by LC–MS, and the collected fractions were subsequently separated by CE. Unambiguous identification of the electrophoretic peaks was achieved by injecting the collected fractions separately and spiking the leuprolide crude mixture. Furthermore, structural information about the components of the mixture provided by several semi-empirical migration models has been used to check the accuracy of the structures previously proposed by LC–MS. Combination of the two orthogonal techniques results in an enhancement of their individual selectivity characteristics. © 2002 Published by Elsevier Science B.V.

Keywords: Peptides; Leuprolide

1. Introduction

The novel analytical needs of researchers working in pharmaceutical, clinical, environmental and biotechnological sciences often require the analysis of complex mixtures of structurally-related compounds. Moreover, in many cases some of the analytes may be present at very low concentration [1,2]. The development of efficient, selective and sensitive methods is necessary to separate and characterize these complex samples. However, single separation methods often lack the peak capacity required for a complete separation and quantitation of all the mixture components and, as a consequence, complementary separation techniques must be employed [3–7]. The application of multidimensional separation approaches, with enhanced separation efficiency, may permit the unambiguous characterization of complex mixtures. Moreover, with a suitable technical implementation, fully-automated, high-throughput multidimensional separation systems are possible [5–7].

☆Presented at the 30th Scientific Meeting of the Spanish Group of Chromatography and Related Techniques/1st Meeting of the Spanish Society of Chromatography and Related Techniques, Valencia, 18–20 April, 2001.

*Corresponding author. Tel.: +34-9-3402-1276; fax: +34-9-3402-1233.

E-mail address: vsanz@apolo.qui.ub.es (V. Sanz-Nebot).

0021-9673/02/$ – see front matter
PII: S0021-9673(02)00025-0

Several multidimensional separation approaches have been described in the literature for the analysis of complex protein or peptide mixtures [3–7]. Two-dimensional (2D) gel electrophoresis with off-line mass spectrometry has been the method traditionally employed in this field [4]. However, this technique often results in slow, non-reproducible and labour intensive analysis. Reversed-phase liquid chromatography (LC) using ionic-pair reagents, such as trifluoroacetic acid, in the mobile-phase, has been widely used for the separation of peptides and proteins [8,9]. Furthermore, LC coupled to mass spectrometry (MS) leads to fast and unambiguous characterization of complex mixtures [7,10,11]. A two-dimensional LC system with tandem mass spectrometric detection (LC–LC–MS–MS) has been recently presented as an alternative approach for structural proteomics [7]. Although these two-dimensional systems are widely accepted because they provide valuable information, the mechanisms of the coupled separation techniques are based on the same principles. It is expected that the more orthogonal the mechanisms of the coupled separation techniques are, the better the combined resolving power.

Capillary electrophoresis (CE) has demonstrated its ability to perform efficient and selective separation and characterization of peptides and protein mixtures [12–14]. The mechanism of CE separations based on the charge-to-mass ratio can be considered orthogonal to the partitioning mechanism of LC separations, which is mainly based on the hydrophobic character of the molecules. Combined separation approaches exploiting this orthogonality can be advantageous to obtain separations with enhanced selectivities.

The off-line combination of LC with CE has been extensively explored to perform tryptic mapping of certain proteins [15–21]. The tryptic digests are firstly manually fractionated, according to hydrophobicity, by LC, and the components of the fractions are then separated, according to their charge-to-mass ratio, by CE. Moreover, LC fractions can be easily concentrated prior to the second separation step, which leads to better limits of detection. The characterization capability of these LC–CE systems is increased when on-line MS detection is incorporated in any of the dimensions [20]. On-line or ‘comprehensive’ LC–CE instruments have also been developed [5,22–24]. In these systems the continuous sampling of the chromatographic eluent and its sequential analysis by CE represent an important throughput improvement as compared with the fraction collection mode. However, reported limits of detection are rather poor, because the sample is diluted in the first chromatographic dimension. To overcome this problem, micro-LC–CE systems with laser-induced fluorescence (LIF) detection are being explored for the analysis of diluted analytes [23,24]. As part of these attempts to develop high-throughput and sensitive methodologies for peptide and protein multidimensional analysis, a few microchip devices have been recently described [25], that are able to perform ‘comprehensive’ two-dimensional open tubular capillary electrochromatography and CE (OTCEC–CE), but these laboratory-on-a-chip technologies are still in their early development [26].

The study of the migration behaviour of peptides and proteins in CE may also provide valuable structural information. For example, several semi-empirical approaches have been proposed to relate electrophoretic mobility to the structural parameters of peptides and proteins. These semiempirical models may be very useful to predict electrophoretic mobility and to perform separation optimizations [27,28]. They also offer a valuable characterization tool, particularly useful when on-line MS detection is not available. Several CE–MS interfaces have been described [29], but only a few CE–MS instruments are commercially available [30]. Several years will pass before CE–MS becomes widely used in analytical laboratories.

Solid-phase peptide synthesis (SPPS) procedures are routinely used to obtain new peptide-based drugs. However, these synthetic procedures often result in complex mixtures [31] that need further separation and characterization [11,32]. LC and CE have proved their suitability for the analysis of multicomponent peptide mixtures, but they are often unable to resolve and identify all the components [11,32,33]. In this work, a combination of LC–MS and CE has been used for the separation and characterization of a complex crude mixture obtained from the synthesis of leuprolide, a widely used therapeutic hormone, structurally analogous to the luteinizing hormone-releasing hormone (LHRH). Crude leuprolide was first fractionated and characterized by LC and LC–

MS, then it was analyzed by CE. In order to obtain an unambiguous identification of the electrophoretic peaks, it was necessary to first inject individual fractions previously collected by LC and then the leuprolide crude spiked with these fractions. The combination of the two orthogonal techniques results in an enhancement of their individual selectivity characteristics. Additionally, several semi-empirical approaches relating the electrophoretic mobilities and structural parameters of peptides and proteins have been tested in order to confirm identity assignments made by LC–MS and CE, according to molecular mass (M_r) and to charge-to-mass ratio, respectively.

2. Experimental

2.1. Chemicals and reagents

All the chemicals used in the preparation of buffers and solutions were analytical reagent grade. Crude leuprolide (Pyr–His–Trp–Ser–Tyr–D-Leu–Leu–Arg–Pro–NHEt) used in this study was supplied by Lipotec (Barcelona, Spain). Leuprolide crude was dissolved in water at a concentration of 1 mg/ml, and was stored in a freezer at 0 °C when not in use. Water, with a conductivity lower than 0.045 mS/cm, was obtained by a Milli-Q water purification system (Millipore, Molsheim, France). Acetonitrile (MeCN), trifluoroacetic acid (TFA), formic acid 90%, ammonia solution 25% and acetone were supplied by Merck (Darmstadt, Germany). All samples and eluents were passed through 0.22 μm nylon filters (MSI, Westboro, MA, USA).

2.2. Apparatus

For the LC experiments an ISCO (Lincoln, NE, USA) Model 2350 chromatographic pump with a 100-μl sample loop in a Valco injection valve and a variable-wavelength V^4 absorbance detector (ISCO) operating at 220 nm was used. The chromatographic system was controlled by Chemresearch Chromatographic Data Management System Controller Software (ISCO) running on a personal computer. A 5 μm Kromasil C_8 column (250×4.6 mm I.D.) (BC Aplicaciones Analíticas, Barcelona, Spain) was used at room temperature. An LKB (Bromma, Sweden) time-programmable fraction collector was used to isolate LC fractions.

CE experiments were performed at 25 °C on a P/ACE System 5500 (Beckman Instruments, Palo Alto, CA, USA) equipped with a photodiode array detector (operating between 190 and 230 nm). An untreated fused-silica capillary of 57 cm (50 cm to the detector)×75 μm I.D., purchased from Polymicro Technologies (Phoenix, AZ, USA), was used. Samples were hydrodynamically injected for 3 s at 0.5 p.s.i. (1 p.s.i.=6894.76 Pa). The working voltage was 15 kV. Data were recorded on a personal computer-based data system (P/ACE Station 1.0 interfaced to the Gold System) supplied by Beckman.

2.3. Procedures

2.3.1. LC procedure

Mobile phase composition was optimised at 1 ml/min as described previously [11]. The solutions tested as mobile phases consisted of different MeCN–water mixtures, ranging from 25% to 35% MeCN (v/v) containing a 0.1% (v/v) of TFA. Two of them, containing 31% and 35% (v/v) MeCN, respectively, were selected to fractionate crude leuprolide.

Before beginning fraction collection, three LC–UV analyses were made to check the reproducibility of the separation profile and retention times. In these experiments, a large sample (100 μl of 1 mg/ml leuprolide crude) was injected. The LC column was then directly connected to the fraction collector with a PTFE tube of the same length and diameter as the tube used to connect it to the UV detector cell. Fractions corresponding to single peaks, previously characterized by LC–electrospray ionization (ESI)–MS [11], were automatically collected. The fractions from 10 separate runs were pooled together and lyophilized in plastic vials. This process was repeated in order to obtain two sets of samples; the lyophilisate fractions from the first set were later reconstituted with 50–75 μl of water (except from that corresponding to leuprolide which was redissolved with 375 μl of water) and those from the second set were used to spike equivalent volumes of a 1 mg/ml leuprolide crude mixture. The two sets of fractions will be called standard and spiked set, respectively.

2.3.2. CE procedure

The influence of pH on CE separations was studied using a 75 m*M* formic acid running buffer adjusted to the appropriate pH with 25% ammonia solution. The capillary was rinsed each day, before starting the analysis, with 0.1 *M* aqueous NaOH solution (10 min), water (20 min) and running buffer (30 min) and, finally a 15 kV voltage was applied for 15 min at 25 °C. The capillary was also flushed with buffer for 3 min before each sample injection and was stored overnight filled with running solution.

Peak identification in the electropherogram of leuprolide crude mixture was performed by injecting 1 mg/ml leuprolide crude, as well as both LC fraction sets, using 75 m*M* formic acid at pH 4. A 1 mg/ml solution of leuprolide crude containing 3% (v/v) acetone as electroosmotic flow (EOF) marker was also injected under the same conditions to obtain effective electrophoretic mobilities (m_e) [14]. m_e was calculated as the difference between the apparent mobility of each peptide and the mobility of acetone [14]. Each m_e value was obtained as the average of three replicates.

3. Results and discussion

3.1. CE analysis. Influence of pH of the running electrolyte

Highly efficient and selective peptide separations have been achieved by CE under acidic conditions in unmodified bare fused-silica capillaries [12–14]. Leuprolide crude mixture was analysed using a 75 m*M* formic acid buffer adjusted to different pH values within the range 2–4.5 (Fig. 1). The resulting electropherograms contain a major peak corresponding to the target peptide leuprolide, as well as a number of peaks of unidentified peptide substances.

Except for the peak indicated in Fig. 1, the migration behaviour of the mixture components does not show a great dependence on pH in the studied

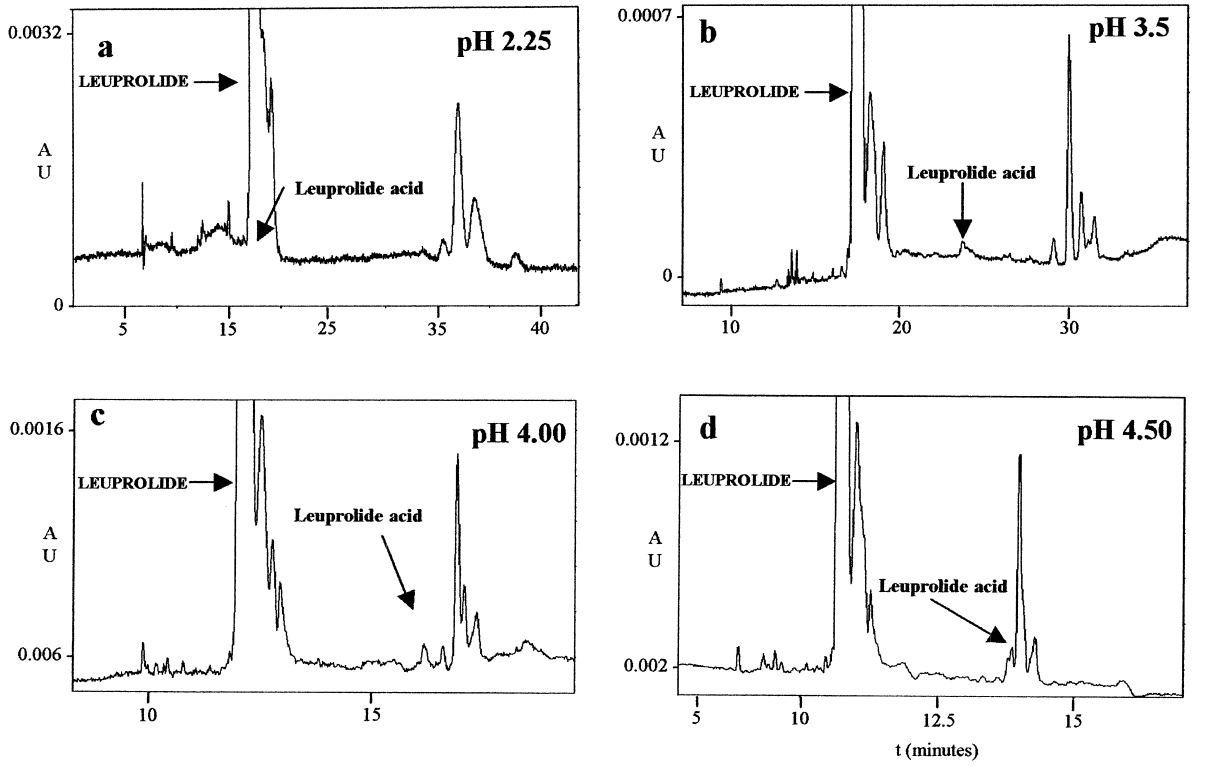

Fig. 1. Effect of buffer pH on leuprolide crude mixture CE separations. pH values (a) 2.25, (b) 3.5, (c) 4.00 and (d) 4.50.

range. This electrophoretic behaviour can be explained by the number and type of ionizable residues present in each peptide. In leuprolide, there are two basic residues—His and Arg—and an acidic phenolic group in Tyr. In a previous work, the p*K* values associated with His and Tyr side chains in triptorelin and buserelin were determined by CE [34]. Experimental p*K* values of Tyr and His in triptorelin and buserelin are about 9.75 and 6.13, respectively. The dissociation of the Arg guanidine group was not determined, but the value given by Lehninger for free Arg is 12.40 [35]. Therefore, the three residues must be fully protonated over the pH range studied and no changes of the peptide net charges are to be expected if only these ionizable groups are present. However, an increase in pH provoked a considerable reduction in the m_e of the peak indicated in Fig. 1, which probably can be attributed to a change in the ionization state. Within the pH range considered, this can only be explained by the deprotonation of a carboxyl group. Lehninger's p*K* values for carboxyl groups range from 2.34, for a free α-COOH, to 4.25, in a Glu residue [35]. In our case, this extra carboxyl group should be in a terminal position, because no glutamic or aspartic acids are expected in leuprolide impurities. In fact, in the last step of the synthesis of leuprolide, leuprolide acid is amidated to give the final leuprolide ethylamide. The presence of leuprolide acid in leuprolide crude mixture is verified by injecting leuprolide crude spiked with a leuprolide acid standard at pH 4 (data not shown). CE at pH 3.5–4 using a 75 m*M* formic acid buffer (Fig. 1b and c) is therefore a useful tool for the simple and rapid monitoring of the final amidation step of leuprolide synthesis. This control cannot be made by LC, because the chromatographic retention of leuprolide acid is not pH-dependent and leuprolide acid co-elutes with leuprolide under the pH conditions studied [11]. Optimum CE separations of the leuprolide mixture components are obtained using a 75 m*M* formic acid buffer at pH 4 (Fig. 1c).

3.2. LC analysis. LC–MS characterization

The separation and characterization of the components of the leuprolide crude mixture was previously performed by LC and LC–MS [11]. The composition of the mobile phase was optimized using the LSER (linear solvation energy relationships) method [8,11,32], and then, characterization by LC–MS was conducted under optimum separation conditions [11]. Typical UV-chromatograms obtained by 31% (v/v) or 35% (v/v) of MeCN mobile phases are shown in Fig. 2a and b, respectively. Masses measured by LC–ESI–MS and structures proposed for labelled peaks in Fig. 2a and b are shown in Table 1. Although characterization was possible by means of LC–ESI–MS [11], the components of the mixture were not fully-resolved. Further separation is therefore desirable and CE can be a useful alternative, as it utilizes a different separation mechanism.

3.3. LC fractionation of leuprolide crude mixture

In order to identify the peaks observed in the electropherogram of leuprolide crude mixture shown in Fig. 1c, labelled chromatographic peaks in Fig. 2a and b were collected [15–21]. A mobile phase with an MeCN percentage of 31% (v/v) was selected to isolate the peaks that elute close to leuprolide (Fig. 2a), while a mobile phase with 35% (v/v) of MeCN permits a rapid collection of the more hydrophobic substances at higher retention times (Fig. 2b). Fractions from several runs were pooled, lyophilized and redissolved in a smaller volume, to concentrate the impurities. A direct identification from the injection of the standard set of fractions is difficult, because migration times are not reproducible enough between successive runs. This could be attributed to slight changes in the electroosmotic flow between different runs, which arises from the reversible modification of the capillary inner wall. Constant electrophoretic mobilities could be obtained by sequential injection of leuprolide crude sample (three runs), but migration times always show small differences that preclude an automatic and unambiguous peak identification in such a complex sample from the fractions of the standard set. Therefore, a second set of spiked samples had to be prepared to identify the peaks in the electropherogram of the leuprolide crude. Nevertheless, co-migration of some of the impurities and their heterogeneous concentrations in leuprolide crude mixture made the information obtained from the injection of the fractions of the standard set very helpful.

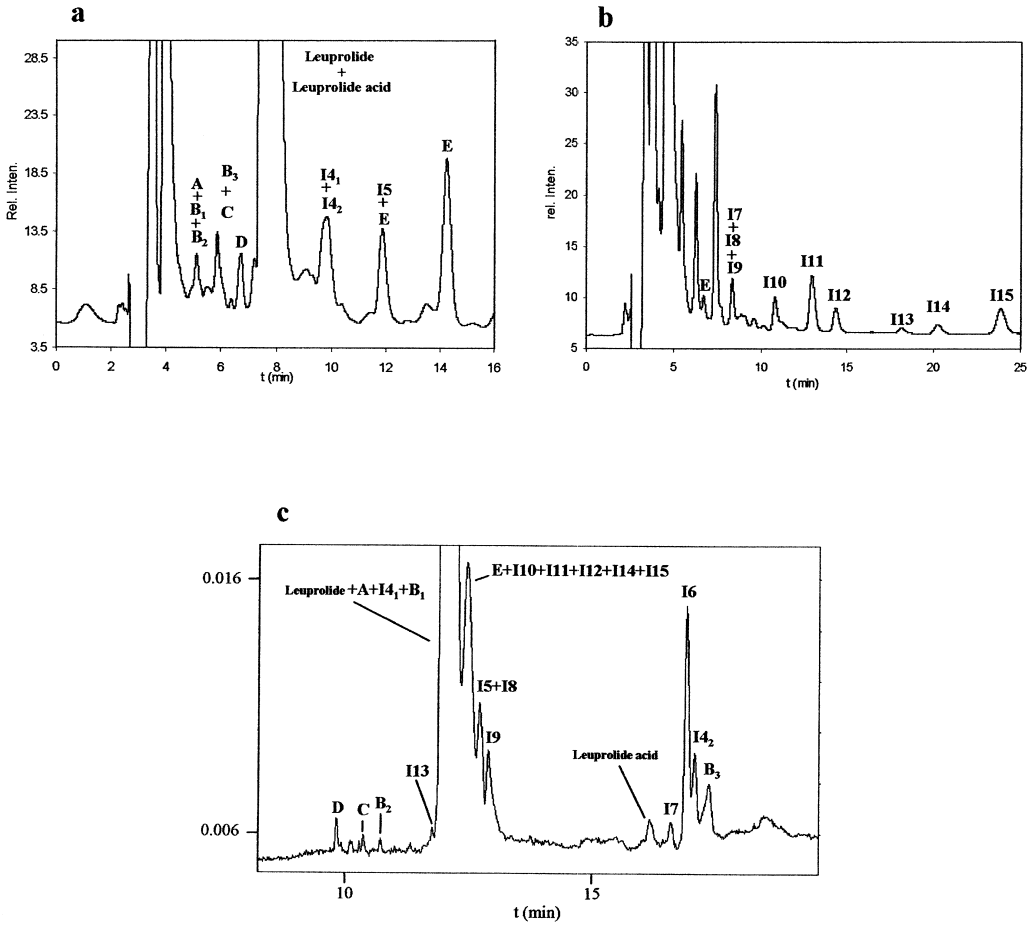

Fig. 2. Comparison of (a and b) LC and (c) CE separations of leuprolide crude mixture. LC mobile phases: (a) water–MeCN (69:31, v/v) with 0.1% (v/v) TFA and (b) water–MeCN (65:35, v/v) with 0.1% (v/v) TFA. CE running buffer: 75 m*M* formic acid, adjusted to pH 4 with NH_3. Other conditions are given in Section 2.

3.4. Identification of CE peaks

Standard and spiked sets of fractions collected by LC were separated under optimal conditions by CE (Figs. 1c and 2c). Peak assignment was made according to these CE results and structures previously proposed by LC–MS [11]. There are great differences in selectivity between the two techniques; for example the elution orders change and some impurities that were previously separated by LC (i.e. *tert*.-butylated impurities) comigrate, as can be seen in Fig. 2c. The partitioning mechanism in LC, mainly based on the hydrophobic character of the analyte, is complementary to the charge-to-mass ratio-based separation mechanism of CE. Moreover, they are considered orthogonal because correlation between capacity factors in LC and mobilities in CE are very low [15]. Thus, impurities $I4_1$ and $I4_2$, that co-elute by LC, can be baseline resolved by CE as can be seen in Fig. 3a (standard set injection of LC fraction $I4_1+I4_2$) and Fig. 3b (spiked set injection of the same fraction). One of the impurities co-migrates with the target peptide, and as a consequence, the leuprolide diastereoisomer $I4_1$ is assigned to the first peak separated by CE. In Fig. 3c the electropherogram corresponding to the leuprolide crude mixture spiked with impurity I13 is shown. Elution time for I13 is more than 10 min higher than the time for

Table 1
LC data obtained from the analysis of leuprolide crude mixture. Proposed structures by LC–MS [11]

Impurity	LC analysis			
	t_R (min)		Measured M_r	Proposed sequence [11]
	31% MeCN	35% MeCN		
A			1211.1	Leuprolide with a reduced tryptophan
B_1	5.11		1211.1	Leuprolide with a reduced tryptophan
B_2			1365.5	Additional arginine insertion
B_3	5.87		1225.0	Leuprolide with an oxidized tryptophan
C			1346.1	Additional histidine insertion
D	6.73		1105.2	Not identified
Leuprolide	7.39		1209.3	[a]Pyr–His–Trp–Ser–Tyr–D-Leu–Leu–Arg–Pro–NH–CH_2–CH_3
Leup. acid			1182.3	Pyr–His–Trp–Ser–Tyr–D-Leu–Leu–Arg–Pro–COOH
$I4_1$	9.87		1209.2	Leuprolide diastereoisomer
$I4_2$			1320.4	[b]Fmoc–His–Trp–Ser–Tyr–D-Leu–Leu–Arg–Pro–NH–CH_2–CH_3
I5	11.85		1335.5	Formation of a substituted guanidine
E		6.80	1265.7	*tert.*-butylated leuprolide
I6	14.20		1124.4	Lack of a serine
I7			1053.3	Lack of an arginine
I8		8.35	1321.9	Doubly *tert.*-butylated leuprolide
I9			1395.5	Additional tryptophan insertion
I10		10.80	1308.8	Not identified
I11		12.96	1265.5	*tert.*-butylated leuprolide
I12		14.34	1265.7	*tert.*-butylated leuprolide
I13		18.35	1088.0	Not identified
I14		20.25	1265.5	*tert.*-butylated leuprolide
I15		23.84	1265.6	*tert.*-butylated leuprolide

[a] Pyr: pyroglutamic acid or 5-oxo-proline.
[b] Fmoc: 9-fluorenylmethoxycarbonyl.

leuprolide (Table 1), but these compounds are not baseline resolved in the electrophoretic conditions used. Fig. 2c shows an electropherogram of leuprolide crude mixture under optimal CE separation conditions. Identified electrophoretic peaks have been labelled according to their original LC denomination. Table 2 summarizes the electrophoretic mobilities (m_e) and identities of the peaks resulting from CE analysis of leuprolide crude mixture.

3.5. Study of the electrophoretic behaviour using semiempirical models

There are several semi-empirical approaches that relate electrophoretic mobility in free solution CE to structural parameters of proteins and peptides [27,28]. These models can be used to predict electrophoretic mobilities but also for optimizing CE separations, studying structural modifications, charge characteristics and conformations [27,28,36].

In general, the electrophoretic mobility (m_e) of a peptide is proportional to its charge (q) and inversely proportional to its Stoke's radius (r). The r is generally expressed in terms of molecular mass (M_r), because the volume of a molecule is proportional to its mass if the density is constant [37]. M_r can be easily determined from the amino acid sequence of the peptide or, as in this case, measured by MS [27,28]. The classical equations describing semi-empirical models are deduced from assumptions concerning the peptide shapes and the forces that they experience during electrophoretic motion [27,28]. The general form of the equation relating electrophoretic mobility, molecular mass, and molecular charge is as follows:

$$m_e = A \cdot \frac{q}{M_r^{\alpha}} \quad (1)$$

where A is a constant, q is the peptide charge and,

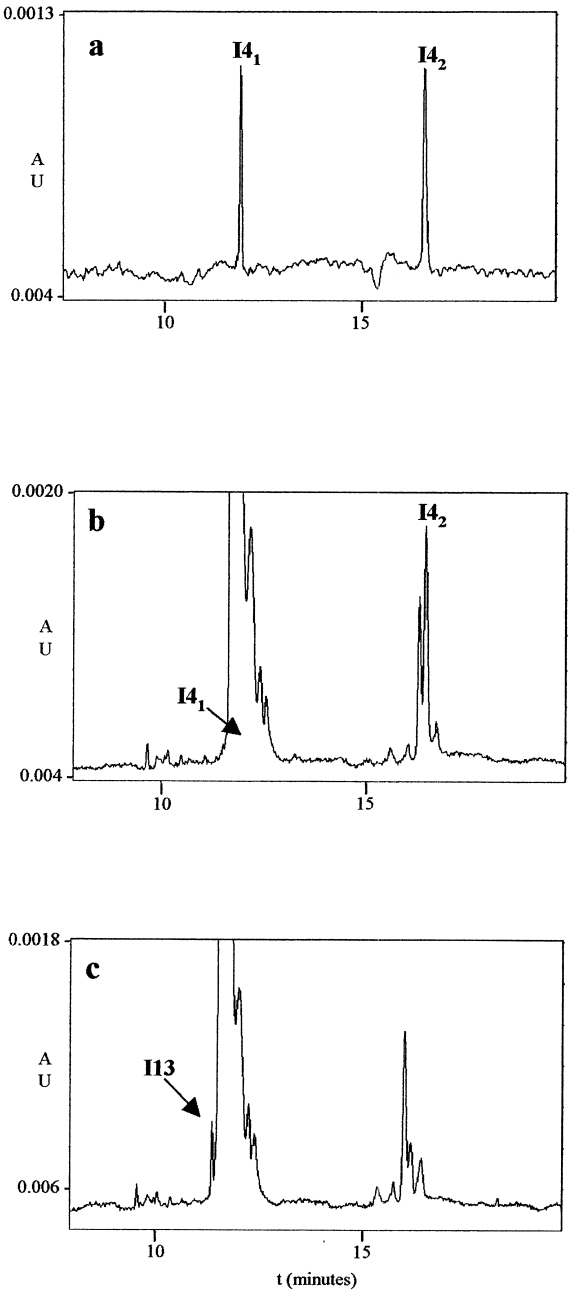

Fig. 3. Electropherograms of several LC fractions corresponding to: (a) I4 (standard set), (b) I4 (spiked set) and (c) I13 (spiked set).

for the various empirical models, $\alpha = 1/3$ for the Stoke's law, $\alpha = 1/2$ for the classical polymer model, and $\alpha = 2/3$ for Offord's surface area law [27,28,37]. In general, for peptides, α approaches 1/3 when peptides are modelled as spherical particles, that have high charge densities; α approaches 1/2 when the peptide is considered as a classical polymer with a lower charge density; and, α approaches 2/3 for larger and more rigid structures, which experience frictional forces that are proportional to the surface area of the molecule during electrophoretic motion.

The molecular charge is an important variable in estimating the electrophoretic mobility. A simple estimation of peptide charge can be obtained at any pH using the peptide pK values and Sillero and Ribeiro expression [38], based on the Henderson–Hasselbach equation:

$$q = \sum_{n=1-4} \frac{P_n}{1 + 10^{\mathrm{pH} - \mathrm{p}K(P_n)}} - \sum_{n=1-5} \frac{N_n}{1 + 10^{-\mathrm{p}K(N_n) - \mathrm{pH}}} \quad (2)$$

where P_n and N_n are the number of cationic (i.e. P_1 = tNH$_2$, P_2 = His, P_3 = Arg and P_4 = Lys) and anionic (i.e. N_1 = tCOOH, N_2 = Asp, N_3 = Glu, N_4 = Cys and N_5 = Tyr) amino acid residues, respectively, and p$K(P_n)$ and p$K(N_n)$ are the negative logarithms of the ionization constants of these amino acids. The accuracy of the charge calculated in this way directly depends on the reliability of the pK_a values employed. In this work, Lehninger's pK values for the individual amino acid residues have been used, instead of the real pK values, which were not available [35]. This fact does not affect the accuracy of the charge calculation because pK values of His, Tyr and Arg are very high, and consequently, these residues are fully protonated at the experimental conditions used. The calculated charge values for the studied peptides at pH 4 are shown in Table 2.

3.5.1. Confirmation of structures proposed by LC–MS

Proposed structures for leuprolide crude mixture components, based on experimental M_r values obtained by LC–MS [11], can be confirmed by studying the structural information provided by the CE analysis. The simple estimation of charge provided by Eq. (2) was used in conjunction with the three semiempirical models previously mentioned, to correlate m_e and q/M_r^{α} according to Eq. (1). m_e values for leuprolide and its associated side products (Table

Table 2
CE data obtained from the analysis of leuprolide crude mixture. Calculated charge[a]. New proposed structures

Impurity	CE analysis			
	t_m (min)	Mobility (10^{-5} cm^2 V^{-1} s^{-1})	q (pH 4)	New proposed sequences
D	9.617	21.1	Not identified	
C	10.133	19.7	2.98	
B_2	10.450	18.5	2.99	
I13	11.316	16.2	Not identified	
Leuprolide			1.99	
A	11.617	15.5	1.99	
$I4_1$			1.99	
B_1			1.99	
E			1.99	
I10			Not identified	
I11	12.117	14.4	1.99	
I12			1.99	
I14			1.99	
I15			1.99	
I5	12.367	13.8	1.99	
I8			1.99	
I9	12.517	13.5	1.99	
Leup. acid	15.550	8.59	1.01	
I7	15.967	8.06	0.99	
I6	16.250	7.71	1.99[b] 0.99[c]	Lack of a serine and arginine is degradated to citrulline
$I4_2$	16.400	7.54	1.99[b] 1.00[c]	A substituted guanidine loses an amino group
B_3	16.667	7.23	1.99[b] 1.00[c]	Leuprolide with an oxidized histidine

[a] Charge calculation was performed using Sillero and Riberio equation (Eq. (2) [38]) and Lehninger's pK values: His (6.00), Tyr (10.00) and Arg (12.40).
[b] These charges were calculated taking into account the sequence proposed by LC–MS [11].
[c] These charges were calculated by taking into account the new proposed sequences.

2) were plotted versus $q/M_r^{1/3}$, $q/M_r^{1/2}$ and $q/M_r^{2/3}$ in Fig. 4a–c, respectively. According to these representations, there is a linear correlation between m_e and q/M_r^{α}, except for impurities I6, $I4_2$, B3. Charge calculations are based on the amino acidic sequence proposed after separating and measuring M_r by LC–MS. An erroneous sequence assignment of a mass measured by LC–MS can lead to an erroneous charge estimation. This is particularly important when ionisable residues are involved. Thus, at pH 4, the migration behaviour of I6, $I4_2$ and B3 does not agree with the doubly-charged structures first proposed. In order to obtain good correlations between their m_e and their charge-to-mass ratio, they must be single charged at pH 4. New sequences for these impurities are proposed according to the observed migration behaviour. I6 was first interpreted as lacking a serine with respect to the sequence of leuprolide. The new structure proposed for I6 also suggests the degradation of arginine to citrulline, which is a non-ionizable residue with the same mass as arginine [32]. $I4_2$ has been related with I5 and it has been attributed to a substituted guanidine which has lost an amino group, and as a consequence is singly charged at pH 4. On the other hand, the oxidation of tryptophan in leuprolide was originally proposed to explain the molecular mass of B3. However, according to CE data, an ionizable residue must be oxidized instead of tryptophan. Oxidation of histidine is also common during SPPS procedures [31]. Table 2 shows the new proposed sequences and calculated charges for I6, $I4_2$, B3. This new set of proposed structures, with known charges and masses, is used to recalculate correlations using the semiempirical models (Eq. (1)). Fig. 4d–f show the new linear correlations between m_e and $q/M_r^{1/3}$, $q/M_r^{1/2}$ and $q/M_r^{2/3}$ for leuprolide and its associated side products. Correlation coefficients (r^2) are given in

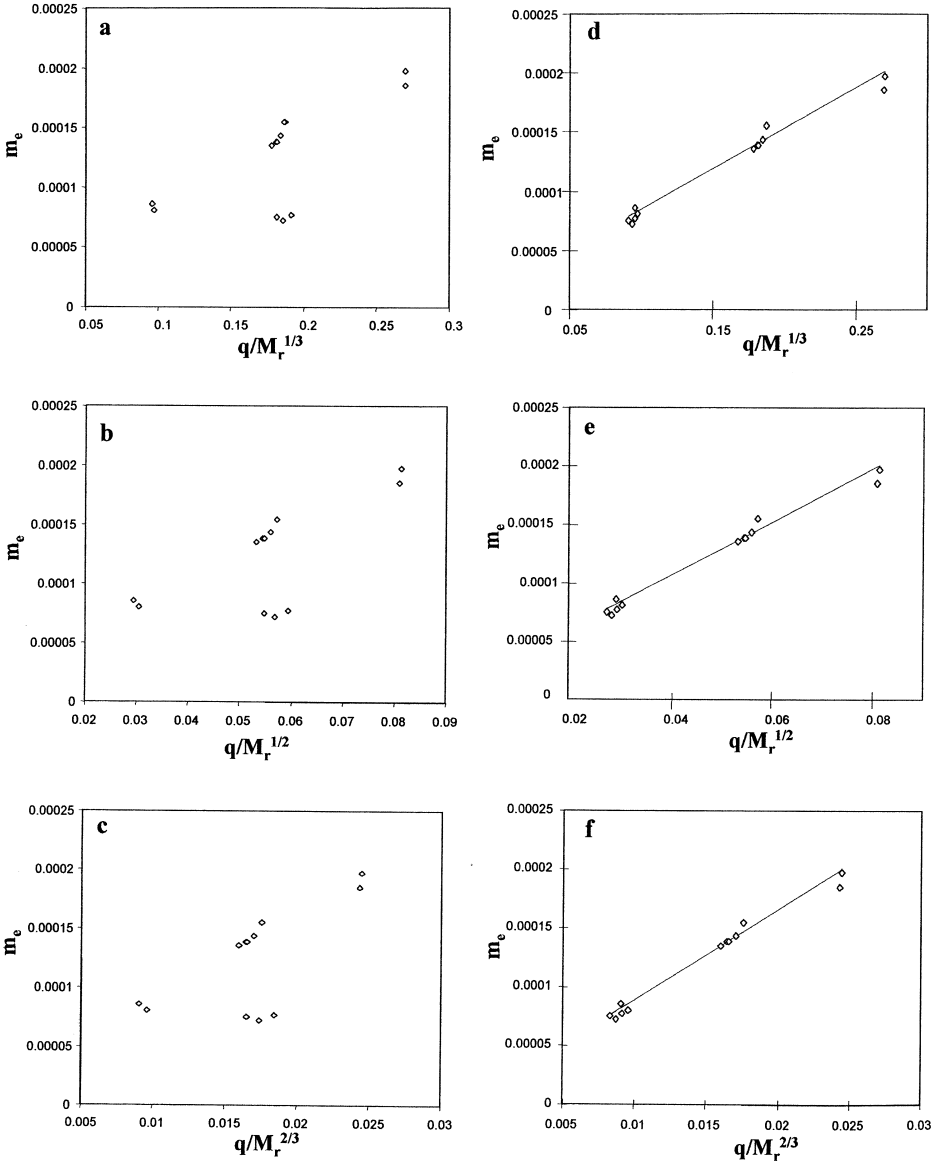

Fig. 4. Correlation of electrophoretic mobility (m_e), before and after proposing new structures, with (a and d) $q/M_r^{1/3}$ (Stoke's law), (b and e) $q/M_r^{1/2}$ (classical polymer equation) and (c and f) $q/M_r^{2/3}$ (Offord's surface area). The equation parameters and the correlation coefficients for the linear least squares fit are given in Table 3.

Table 3. The good correlations observed confirm the validity of the new proposed structures.

3.5.2. Testing semiempirical models for peptide migration behaviour

As can be observed from a comparison of the correlation coefficients given in Table 3, a slightly better linear correlation is observed when Offord's surface area model is applied. This suggests rigid peptide structures which undergo electrophoretic motion and experience frictional forces proportional to the surface area of the molecule [39]. Janini et al. already showed that these three models will yield comparable correlations for subsets of peptides with

Table 3
Summary of linear least squares parameters obtained from application of classical semi-empirical models in leuprolide crude mixture, uncorrected and corrected for charge suppression

Linear parameters $m_e = a + bX$	Stoke's law		Polymer law		Offord's law		Grossman's law
	$q/M_r^{1/3}$	$\ln(1+q)/M_r^{1/3}$	$q/M_r^{1/2}$	$\ln(1+q)/M_r^{1/2}$	$q/M_r^{2/3}$	$\ln(1+q)/M_r^{2/3}$	$\ln(1+q)/n^{0.435}$
a	1.715×10^{-5}	-4.741×10^{-5}	1.449×10^{-5}	-5.288×10^{-5}	1.187×10^{-5}	-5.795×10^{-5}	-4.679×10^{-5}
b	6.826×10^{-4}	1.910×10^{-3}	2.297×10^{-3}	6.470×10^{-3}	7.723×10^{-3}	2.184×10^{-2}	4.614×10^{-4}
r^2	0.9653	0.9848	0.9716	0.9880	0.9766	0.9873	0.9723

narrow molar mass ranges, as is our case. Nevertheless, it was also observed that better correlations were obtained when Offord's model was applied to a large set of peptides with a wide molar mass range [39].

Charge calculation using Sillero and Ribeiro's equation (Eq. (2)) is likely to result in an overestimation of q under conditions where the protein has an appreciable charge [27,28]. As the total charge on the peptide increases, the effect of other additional charges on its mobility decreases, resulting in an 'ineffectiveness' of a part of the peptide charge (charge suppression effect). The semi-empirical models described above can be improved if the electrostatic charge suppression effect is considered in Eq. (1). This means substituting the direct proportionality by a logarithmic dependence. Thus, Eq. (1) can be re-written as Eq. (3):

$$m_e = A \cdot \frac{\ln(1+q)}{M_r^{\alpha}} \quad (3)$$

This simple compensation for charge suppression is inadequate for proteins, where the magnitude of charge suppression is greater and the mechanisms are more complex [27,28]. Cifuentes and Poppe considered that an adjustable curvature better describes the charge suppression effect when charge increases, and they proposed a model corrected as $\ln(1+Bq)$ [40]. Fig. 5a–c shows plots of m_e versus $\ln(1+q)/M_r^{1/3}$, $\ln(1+q)/M_r^{1/2}$ and $\ln(1+q)/M_r^{2/3}$, respectively. Correlation coefficients (r^2) are given in Table 3. Linear correlation coefficients r^2 are slightly better when models corrected for charge suppression are used (Table 3).

3.5.3. Characterization of impurities unidentified by LC–MS

m_e can be estimated for impurities with known molecular mass using the improved semi-empiricals models. Conversely, molecular charge can be estimated for the unidentified impurities whose M_r has been previously measured by MS. Table 4 shows calculated charge values for impurities D, I13 and I10. At this point Grossman's semi-empirical relationship [27,28,41] can be introduced, where the number of amino acid residues of the peptide structure (n) is taken into account:

$$m_e = A \cdot \frac{\ln(1+q)}{n^{0.43}} \quad (4)$$

where A is again an adjustable constant. A good linear correlation between m_e and $\ln(1+q)/n^{0.43}$ is shown in Fig. 6. Better correlations are observed using models that consider molecular mass (Table 3), but Grossman's relationship is useful to estimate the number of amino acid residues in the unidentified D, I10 and I13 impurities. Table 4 shows charge values obtained for D, I10 and I13, using the classical models corrected for charge suppression, and their number of amino acids calculated from Grossman's equation. The information obtained about the ionizable groups, estimating the charge with the first type of models, and the number of amino acid residues calculated with Grossman's relation, can be used to propose a sequence for D, I10 and I13. This completes the characterization of leuprolide crude mixture. According to this information, D would have a charge of about 3 and contain 8 amino acid residues, which agrees with an additional His insertion in a leuprolide structure without the Pro and Tyr residues. For I10, a leuprolide with an additional Pro residue agrees with a doubly charged decapeptide. However, further information is needed to propose a structure for the doubly charged I13 octapeptide. An unambiguous

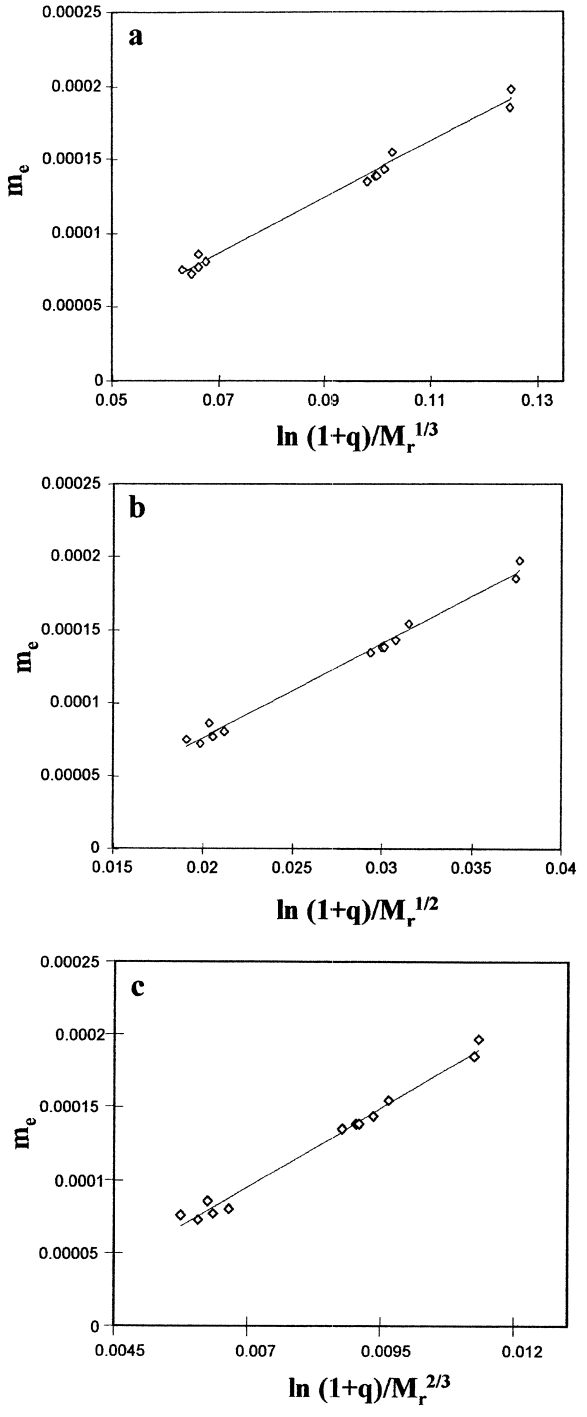

Fig. 5. Charge suppression effect is taken into account in plots of m_e correlations with (a) $\ln (1+q)/M_r^{1/3}$, (b) $\ln (1+q)/M_r^{1/2}$ and (c) $\ln (1+q)/M_r^{2/3}$. The equation parameters and the correlation coefficients for the linear least squares fit are given in Table 3.

Table 4
Calculated charge and amino acid residue numbers for the unidentified impurities

Stoke's law		D	I13	I10
Stoke's law	*q*	3.06	2.09	1.99
	n	8.24	8.15	9.40
Polymer's law	*q*	2.89	1.99	2.00
	n	7.68	7.63	9.50
Offord's law	*q*	2.74	1.90	2.02
	n	7.17	7.15	9.60

q is calculated using each classical model corrected for charge suppression. *n* is calculated using Grossman's equation.

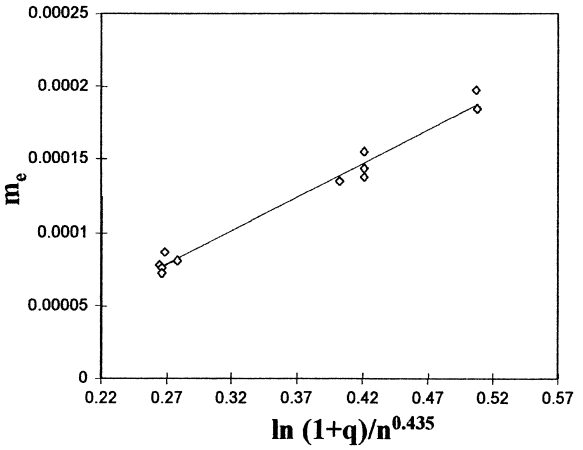

Fig. 6. Correlation of electrophoretic mobility (m_e) with $\ln (1+q)/n^{0.435}$ (Grossman's law). The equation parameters and the correlation coefficient for the linear least squares fit is given in Table 3.

characterization of complex peptide mixtures would be only possible if separated peptides could be sequenced by means of MS–MS. A sequential automated combination of LC–MS–MS and CE–MS–MS may provide maximum separation and characterization capabilities to perform high-throughput analysis of complex, diluted mixtures. The performances of these combined techniques should be extensively explored in the future for the advantage of many people working in proteomic, biotechnological, environmental and pharmaceutical research.

4. Concluding remarks

Combination of LC–MS and CE has been successfully applied for the separation and characterization of a leuprolide crude mixture. The orthogonal separation mechanisms that are involved in both techniques result in complementary separation selectivities. CE analysis of collected LC fractions leads to the identification of the electrophoretic peaks. The applicability of semi-empirical models that describe the migration behaviour in CE has been assessed and the additional structural information obtained has been used to confirm structural assignments made on the basis of molecular mass measured by LC–MS. Furthermore, new sequences have been proposed for several impurities that remained unidentified.

References

[1] A.F.R. Hümer, G.I. Aced, M.D. Perkins, R.N. Gürsoy, D.S.S. Jois, C. Larive, T.J. Siahaan, C. Schöneich, Anal. Chem. 69 (1997) 29R.
[2] A. Karcher, Z. El Rassi, Electrophoresis 20 (1999) 3280.
[3] H.Y. Gao, Y.F. Shen, T.D. Veenstra, R. Harkewicz, G.A. Anderson, J.E. Bruce, L. Pasatolic, R.D. Smith, J. Microcol. Sep. 12 (2000) 383.
[4] P. Davidsson, A. Westman, M. Puchades, C.L. Nilsson, K. Blennow, Anal. Chem. 71 (1999) 642.
[5] A.W. Moore, J.W. Jorgenson, Anal. Chem. 67 (1995) 3448.
[6] H. Yamamoto, T. Manabe, T. Okuyama, J. Chromatogr. 515 (1990) 659.
[7] J. Yates, M. Washburn, D. Woelters, Paper L1601 presented at the 14th International Symposium on Microscale Separations and Analysis, Boston, MA, January, 2001.
[8] C.K. Larive, S.M. Lunte, M.Z. Zhong, M.D. Perkins, G.S. Wilson, G. Gokulrangan, T. Williams, F. Afroz, C. Schöneich, T.S. Derrick, C.R. Middaugh, S. Bogdanowich-Knipp, Anal. Chem. 71 (1999) 389R.
[9] V. Sanz-Nebot, F. Benavente, I. Toro, J. Barbosa, Chromatographia 53 (2001) S167.
[10] P.A. D'Agostino, J.R. Hancock, L.R. Provost, P.D. Semchuk, R.S. Hodges, J. Chromatogr. A 800 (1998) 89.
[11] V. Sanz-Nebot, F. Benavente, J. Barbosa, J. Chromatogr. A 870 (2000) 315.
[12] S.C. Beale, Anal. Chem. 70 (1998) 279R.
[13] V. Kasicka, Electrophoresis 20 (1999) 3084.
[14] V. Sanz-Nebot, F. Benavente, I. Toro, J. Barbosa, J. Chromatogr. A 921 (2001) 69.
[15] M. Castagnola, L. Cassiano, R. Rabino, D.V. Rossetti, F.A. Bassi, J. Chromatogr. 572 (1991) 51.
[16] J. Frenz, S.L. Wu, W.S. Hancock, J. Chromatogr. 480 (1989) 379.
[17] S.E. Rudnick Jr., V.J. Hilser, G.D. Worosila, J. Chromatogr. A 672 (1994) 219.
[18] W. Voelter, J. Schutz, O.E. Tsitsiloni, A. Weiler, G. Grüber, G. Paulus, S. Stoeva, R. Lehmann, J. Chromatogr. A 807 (1998) 135.
[19] H.J. Issaq, K.C. Chan, G.M. Janini, G.M. Muschik, Electrophoresis 20 (1999) 1533.
[20] H.J. Boss, D.B. Watson, R.S. Rush, Electrophoresis 19 (1998) 2654.
[21] H.J. Issaq, K.C. Chan, C. Liu, Q. Li, Electrophoresis 22 (2001) 1133.
[22] K.C. Lewis, G.J. Opiteck, J.W. Jorgenson, D.M. Sheeley, J. Am. Soc. Mass Spectrom. 8 (1997) 495.
[23] H. Zhang, J.W. Jorgenson, Paper P1302Ta presented at the 14th International Symposium on Microscale Separations and Analysis, Boston, MA, January, 2001.
[24] I. German, R.T. Kennedy, Anal. Chem. 72 (2000) 5365.
[25] R.D. Rocklin, R.S. Ramsey, J.M. Ramsey, Anal. Chem. 72 (2000) 5244.
[26] J.P. Kutter, Trac-Trends Anal. Chem. 19 (2000) 352.
[27] N.J. Adamson, E.C. Reynolds, J. Chromatogr. B 699 (1997) 133.
[28] A. Cifuentes, H. Poppe, Electrophoresis 18 (1997) 2362.
[29] J. Ding, P. Vouros, Anal. Chem. 71 (1999) 378A.
[30] N.A. Guzmán, J. Chromatogr. B 749 (2000) 197.
[31] P.L. Williams, F. Albericio, E. Giralt, Chemical Approaches to the Synthesis of Peptides and Proteins, CRC Press, Boca Raton, FL, 1997.
[32] V. Sanz-Nebot, F. Benavente, A. Castillo, J. Barbosa, J. Chromatogr. A 889 (2000) 119.
[33] S. Ridge, K. Hettiarachchi, J. Chromatogr. A 817 (1998) 215.
[34] V. Sanz-Nebot, F. Benavente, I. Toro, J. Barbosa, Electrophoresis 22 (2001) 4333.
[35] A.L. Lehninger, D.L. Nelson, M.M. Cox, in: 2nd ed, Principles of Biochemistry, Work Publishers, New York, 1993.
[36] B.R. Sitaram, H.H. Keah, M.T.W. Hearn, J. Chromatogr. A 857 (1999) 263.
[37] E.C. Rickard, M.M. Strohl, R.G. Nielsen, Anal. Biochem. 197 (1991) 197.
[38] A. Sillero, J.M. Ribeiro, Anal. Biochem. 179 (1989) 319.
[39] G.M. Janini, C.J. Metral, H.J. Issaq, G.M. Muschik, J. Chromatogr. A 848 (1999) 417.
[40] A. Cifuentes, H. Poppe, J. Chromatogr. A 680 (1994) 321.
[41] P.D. Grossman, K.J. Wilson, G. Petrie, H.H. Lauer, Anal. Biochem. 173 (1988) 265.