ELSEVIER

Journal of Chromatography B, 772 (2002) 35–44

JOURNAL OF CHROMATOGRAPHY B

www.elsevier.com/locate/chromb

Simplification of complex tryptic digests for capillary electrophoresis by affinity selection of histidine-containing peptides with immobilised metal ion affinity chromatography

Ahmad Amini[b,*], Asish Chakraborty[a], Fred E. Regnier[a]

[a]*Department of Chemistry, Purdue University, West Lafayette, IN 47907-1393, USA*
[b]*Medical Products Agency (MPA), Division of Biotechnology, Box 26, Husargatan 8, S-751 03 Uppsala, Sweden*

Received 24 August 2001; received in revised form 10 January 2002; accepted 14 January 2002

Abstract

This paper reports on the selectivity behaviour of tryptic peptides on a Cu^{2+}-loaded immobilised metal ion affinity chromatography (IMAC) support. Ovalbumin was chosen as a model protein for investigation of the selection and separation of histidine-containing peptides by IMAC off-line coupled with capillary electrophoresis and matrix-assisted laser desorption ionisation time-of-flight mass spectrometry (MALDI–TOF). Two of five histidine-containing peptides in addition to some non-histidine-containing peptides from a tryptic digest of ovalbumin were captured by IMAC. To separate and purify the selected peptides, the IMAC sample was analysed by capillary zone electrophoresis (CZE). The sample was not separated by capillary zone electrophoresis, therefore, micellar electrokinetic chromatography (MEKC) using 10–75 m*M* SDS was used. Analysis of IMAC sample by MEKC, using low concentrations of SDS (10 m*M*) was characterised by MALDI–TOF. When using SDS at 75 m*M*, the migration times of reversed-phase fractions of the IMAC sample, were used to identify the peaks. One of the two selected histidine-containing peptides with two histidine residues was identified, analysing the sample by CZE or MEKC.

Keywords: Tryptic digests; Peptides; Histidine

1. Introduction

Genomics and DNA databases have transformed the way we think about protein characterisation. Instead of chemically sequencing proteins, DNA databases are now being used to predict the sequence [1]. Following trypsin digestion of a protein, masses of a few of the peptide fragments, called signature peptides, are being used to identify the parent gene in a DNA database and the protein sequence is derived from the DNA sequence of the gene. Proteins in complex mixtures can even be identified and sequenced in this way with the aid of separation systems to fractionate and simplify the tryptic peptides through affinity selection. The two critical elements in this approach to protein characterisation are: (i) highly selective and rapid separation systems, and (ii) a mass spectrometer capable of analysing mixtures of peptides. Immobilised metal ion affinity chromatography (IMAC) [2], high-performance liquid chromatography, capillary electrophoresis (CE) and matrix-assisted laser desorption ionisation time-

*Corresponding author. Tel.: +46-1817-4619; fax: +46-1854-8566.

E-mail address: ahmad.amini@mpa.se (A. Amini).

1570-0232/02/$ – see front matter
PII: S1570-0232(02)00034-X

of-flight mass spectrometry (MALDI–TOF) are important tools in this approach to protein characterisation.

The presence of the rare amino acid histidine in a protein or peptide is a crucial factor in binding of them with transition metal ions, such as Cu^{2+} and Zn^{2+}, in IMAC [2–4]. IMAC selection of a complex tryptic digest of a biological sample reduces the number of peptides to a limited number, which can be used as signature peptide/s for identification of the original proteins. A signature peptide is a unique peptide for identification of a protein, in respect to its mass, amino acid sequence, retention behaviour on a reversed-phase chromatographic column, electrophoretic mobility, or a combination of the mentioned properties.

In the signature peptide approach to proteomics [5] all proteins in a sample are digested by trypsin first and then affinity selected in order to simplify the complex sample, which may contain more than one million peptides. Some or all of the selected peptides can be used as the signature peptides for identification and quantification of the proteins present in the original sample. The advantage of this approach is that peptides are easier to analyse than proteins. However, the drawback of this strategy is that the digestion increases the complexity of the sample, which must be fractionated by multidimensional chromatography.

CE has the advantage of being a high efficient separation technique that is rapid and consumes samples at nanolitre levels. This is particularly useful when separating peptides in large numbers of fractions derived from affinity or reversed-phase chromatography. Micellar electrokinetic chromatography (MEKC) extends the range of electrophoretic separations to include partitioning into micelles [6]. MEKC [7–9] has been found to be efficient in the separation of peptides by providing alternative selectivity to that of CE alone [7]. It provides unique selectivity by simultaneously selecting on the basis of molecular size and shape, charge, hydrophobicity, and even the tendency to hydrogen bond in some cases [10].

The great advantages of MALDI–TOF [11,12] over electrospray ionisation (ESI) [13] in the analysis of polypeptides are: (i) the general absence of multiply charged states of the analyte, which allows the analysis of more complex samples, and (ii) higher tolerance to impurities, such as salts and buffer additives [14]. However, the major problem with MALDI–TOF is quenching of ionisation. This problem can be resolved by purification of the analyte sample and/or by increasing the degree of fractionation with liquid chromatography or capillary electrophoresis before MALDI analysis, i.e. by reducing the number of components in the sample presented to the mass spectrometer [14]. This paper examines the analysis of the IMAC fractions by CZE and MEKC off-line coupled with MALDI–TOF.

2. Experimental

2.1. Chemicals

Ovalbumin, SDS (lauryl sulphate), dithiothreitol (DTT), *N*-tosyl-L-lysyl chloromethyl ketone (TLCK), TPCK-treated trypsin, 2-mercaptoethanol (2-hydroxyethylmercapten; β-mercaptoethanol), 4-vinyl-pyridine and α-cyano-4-hydroxycinnamic acid were purchased from Sigma (St Louis, MO, USA) and imidazole from ACROS (NJ, USA). Ammonium acetate was obtained from Fisher Scientific (NJ, USA).

2.2. Capillary electrophoresis

Capillary electrophoresis experiments were performed using a BioFocus™ 3000 Capillary Electrophoresis System (BIO RAD, CA, USA), equipped with a UV detector monitoring a wavelength of 200 nm. A fused-silica capillary 38 cm (34 cm effective length)×50 μm I.D. (Polymicro Technologies, Phoenix, AZ, USA) was kept at a constant temperature of 25.0 °C and the applied voltage varied between 2.5 and 30 kV.

2.3. Immobilised metal affinity chromatography (IMAC) and reversed-phase chromatography (RPC)

IMAC was performed employing a copper-loaded POROS MC column, 4.6 mm diameter, (PE-Biosystems). The column was equilibrated with 0.5 *M* sodium chloride dissolved in 20 m*M* sodium phos-

Table 1
Peptide map over free solution tryptic digest of ovalbumin

Peptide sequence	Mass
ILELPFASGTMSMLVLLPDEVSGLEQLESIINFEK	3862.011
YNLTSVLMAMGITDVFSSSANLSGISSAESLK	3293.628
VHHANENENIFYCPIAIMSALAMVYLGAK	2976.493
NVLQPSSVDSQTAMVLVNAIVFK	2460.317
LPGFGDSIEAQCGTSVNVHSSLR	2374.146
VTEQESKPVQMMYQIGLFR	2284.146
DILNQITKPNDVYSFSLASR	2281.182
EVVGSAEAGVDAASVSEEFR	2008.946
ELINSWVESQTNGIIR	1858.966
ISQAVHAAHAEINEAGR	1773.899
GSIGAASMEFCFDVFK	1708.771
GGLEPINFQTAADQAR	1687.840
LTEWTSSNVMEER	1581.721
YPILPEYLQCVK	1465.776
HIATNAVLFFGR	1345.737
DEDTQAMPER	1209.520
ADHPFLFCIK	1190.603
VASMASEK	822.403
LYAEER	780.389
VYLPR	647.388
GLWEK	632.340
TQINK	603.346
ELYR	580.309
MEEK	536.230

Digestion conditions are given in the Experimental section.

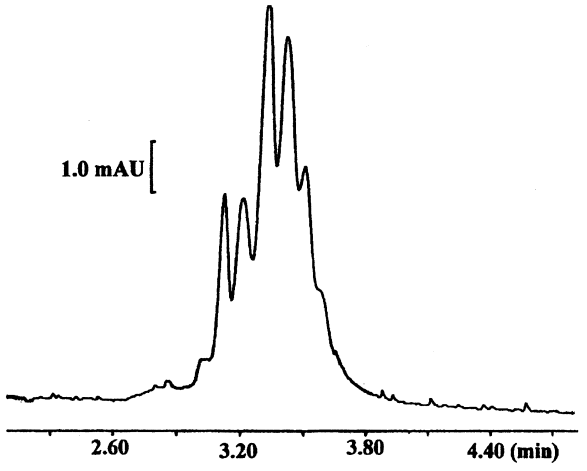

Fig. 1. Analysis of ovalbumin by CZE. Operating conditions: separation capillary, 38 cm×50 μm (effective length 34 cm) thermostated at 25 °C; BGE was 20 m*M* ammonium acetate (pH 7); UV detection at 200 nm; applied voltage was +20 kV. Ovalbumin was dissolved in water at a concentration of 10 μg/ml.

phate buffer at pH 7.0. The elution of the peptides was achieved using a linear salt gradient for 90 min, i.e. 1–20 m*M* imidazole dissolved in the starting buffer. Detection was carried out by UV at 280 nm.

RPC was performed using a PapMap C_{18}™ 2.1× 250 mm Silica C_{18} (PE-Biosystems). An 80-min gradient from 10% ACN to 95% ACN in 0.1% TFA was used to elute peptides from the reversed-phase column at room temperature. Detection was carried out by UV at 215 nm.

All the buffers and solutions were prepared in distilled water, degassed and filtered through a 0.2-μm filter prior to use.

2.4. Proteolysis

Ovalbumin (5 mg) was reduced and alkylated in 1 ml of 0.2 *M* Tris buffer (pH 8.5) containing 8 *M* urea and 10 m*M* dithiothreitol (DTT). The mixture was incubated at 37 °C for 2 h. Vinyl-pyridine as an alkylating agent was added to a final concentration of 20 m*M* and incubated for 2 h. β-Mercaptoethanol was then added to the reaction mixture at a concentration of 40 m*M* in order to interact with excess vinyl-pyridine. The reaction was allowed to proceed at room temperature for 30 min; 0.2 *M* Tris buffer was added to the mixture to reduce the urea concentration from 8 to 3 *M*. Finally, 100 μg TPCK-treated trypsin (2% of the enzyme by weight to that of the protein) was added and incubated for 24 h at 37 °C. The reaction was stopped by adding TLCK in slight molar excess to that of trypsin. The tryptic peptides of the complete digestion of ovalbumin are shown in Table 1, the protein has been sequenced by Nibset et al. [15].

2.5. MALDI–TOF-MS and sample preparation

MALDI analyses were performed on a Voyager DE-RP BioSpectrometry workstation (PE Biosystems) of reflector type time-of-flight mass spectrometer, equipped with a pulsed nitrogen laser working at 337 nm radiation.

The instrument was operated in positive-ion and reflector mode with delayed extraction at an accelerating voltage of 20 kV. Samples were prepared by mixing a 1-μl aliquot with 1-μl matrix solution. A 1-μl aliquot of the mixture was spotted into a well

of the MALDI sample plate and allowed to air-dry before being placed in the mass spectrometer. The matrix was a saturated solution of α-cyano-4-hydroxycinnamic acid in 3% TFA dissolved in ACN/water (50:50).

3. Results and discussion

3.1. Analysis of trypsin digested ovalbumin

Ovalbumin is a conjugated protein, which contains substantial post translational heterogeneity, i.e. glycosylation and phosphorylation. The carbohydrate moiety that constitutes 3% of the molecular mass makes the protein like other glycoproteins, to be present in several isoforms, whereby the character of the carbohydrate moiety in the molecule differs [16–18]. In our study, the presence of at least five different isoforms is revealed, analysing the protein by CE (Fig. 1).

A trypsin digest of ovalbumin was analysed by CE under different conditions. As Fig. 2 illustrates, the separation of tryptic sample is better at pH 8.0 (Fig. 2a) than at pH 9.0 (Fig. 2b) and 7.0 (Fig. 2c). The tryptic sample was also analysed with RPC using a

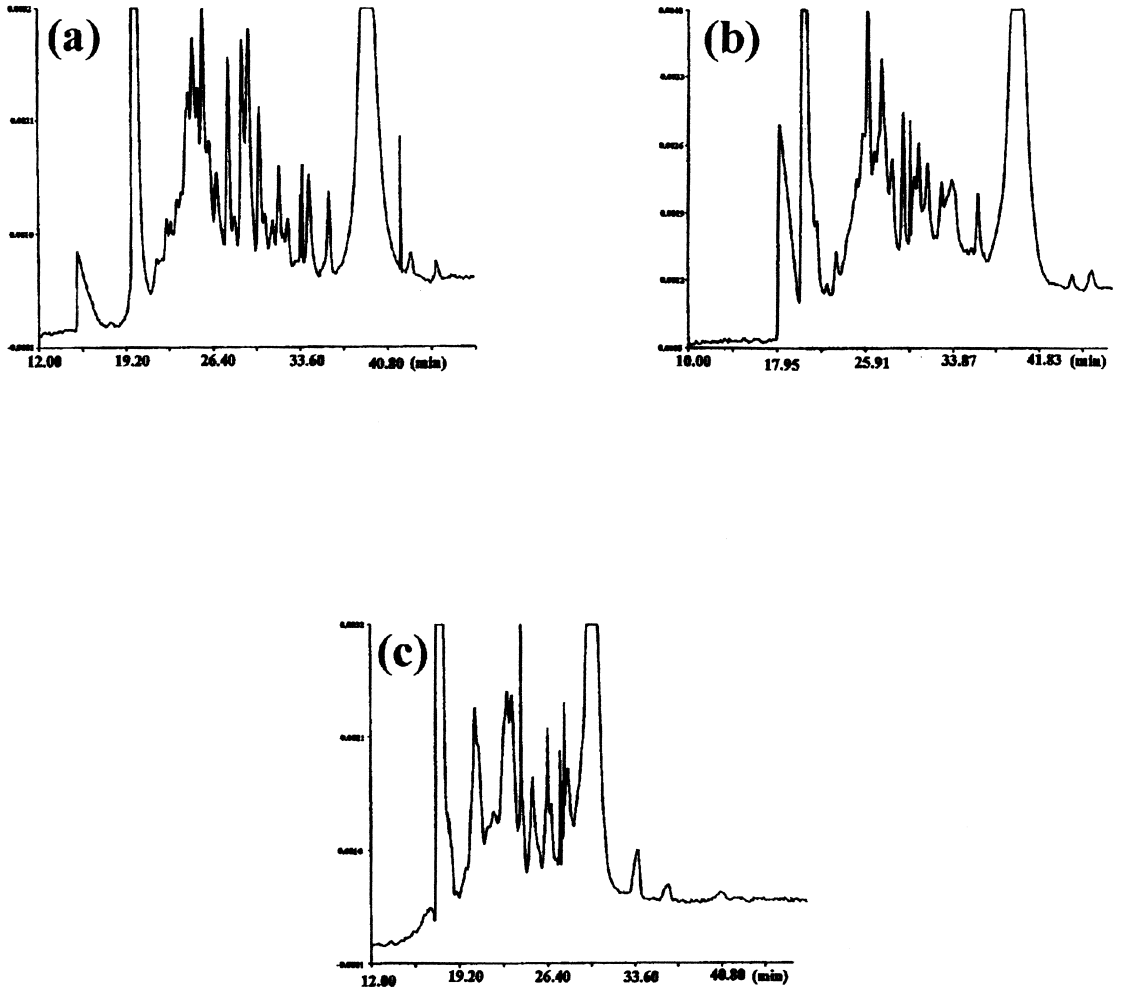

Fig. 2. Separation of trypsin digested ovalbumin by CZE. Operating conditions: BGE was (a) 20 m*M* phosphate buffer (pH 8), (b) 20 m*M* phosphate buffer (pH 9) and (c) 20 m*M* Tris buffer (pH 7); applied voltage was +2.5 kV. Other conditions as in Fig. 1.

linear acetonitrile gradient (Fig. 3). Comparison between the performed separations by CE and RPC showed that RPC detected about 30% more peaks than CE, indicating that separation of the tryptic peptides on the basis of hydrophobicity, being proportional to the size and amino acid residue, is more powerful than electrophoretic separation.

The number of peaks present in Fig. 3 is more than the theoretical number of peptides given in Table 1. The reason may be the presence of digestion additives, such as DTT, vinyl-pyridine, pyridine modified β-mercaptoethanol and trypsin, intact ovalbumin, impurities and miscleaved peptides in the tryptic sample.

MALDI–TOF analysis of the tryptic sample is given in Table 2. Two histidine-containing peptides were identified; however, the concentration of one of the peptides (HIATNAVLFFGR, M_w=1345.737) was very low, resulting in a mass spectrum with low intensity upon MALDI analysis. Other histidine-containing peptides and some of the peptides listed in Table 1 were not detected, perhaps owing to miscleavage and quenching of peptide ionisation in

Table 2
MALDI analysis of tryptic digest of ovalbumin

Peptide sequence	Mass
NVLQPSSVDSQTAMVLVNAIVFK	2460.317
VTEQESKPVQMMYQIGLFR	2284.146
DILNQITKPNDVYSFSLASR	2281.182
EVVGSAEAGVDAASVSEEFR	2008.946
ELINSWVESQTNGIIR	1858.966
ISQAVHAAHAEINEAGR	1773.899
GSIGAASMEFCFDVFK	1708.771
GGLEPINFQTAADQAR	1687.840
LTEWTSSNVMEER	1581.721
YPILPEYLQCVK	1465.776
HIATNAVLFFGR	1345.737
DEDTQAMPER	1209.520

Peptides with a mass less than 1000 and peptides not matching the calculated mass have been excluded.

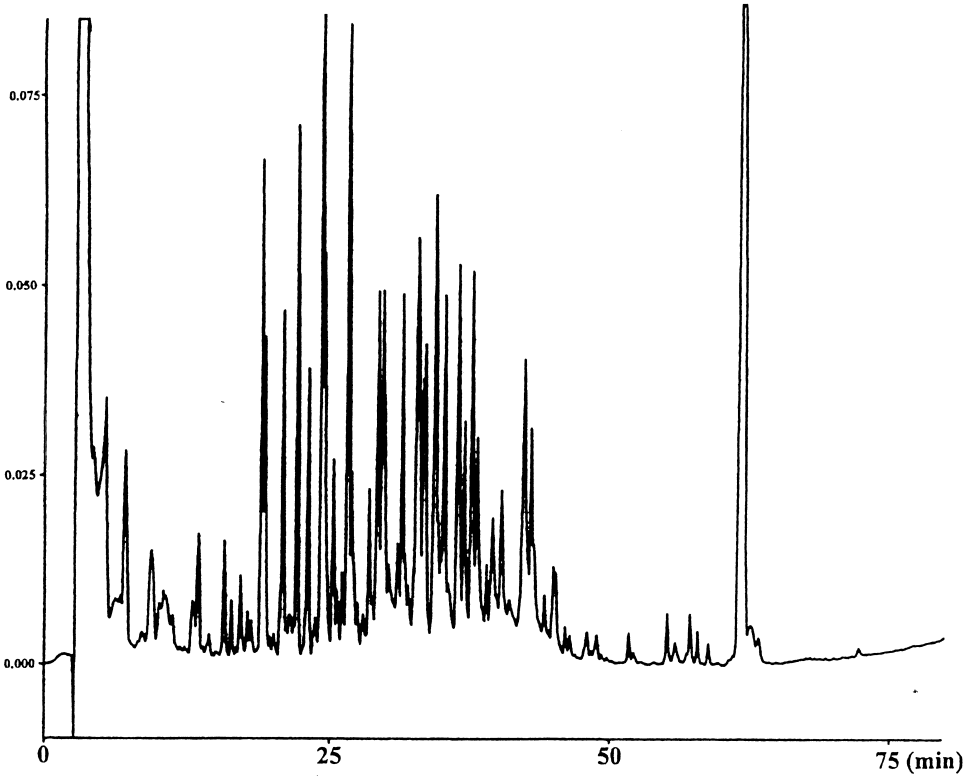

Fig. 3. Reversed-phase chromatography of a tryptic digest of ovalbumin. Conditions: See the Experimental section.

MALDI, because of the complexity of the tryptic sample. In order to reduce the complexity of the tryptic sample, it was analysed by a Cu^{2+}-loaded IMAC support.

3.2. Affinity selection of histidine-containing peptides by IMAC

IMAC is an established technique for selection of histidine-containing peptides [2]. The interaction between the analyte and the stationary phase occurs by chelate complexation with immobilized metal ion, and hydrophobic and electrostatic interactions. However, chelate complexation is the most specific and selective interaction responsible for the affinity selection of peptides, containing histidine, cystein or tryptophane [19].

During the course of trypsin digestion, peptide bonds are cleaved on the C-terminal side of the basic amino acids, producing peptides with an N-terminal amino group, which interacts with the immobilized metal ions [20–24].

The interaction of these peptides with the IMAC-absorbent can be controlled in several ways. Sulfhydryl alkylation of cystein during the digestion process eliminates cystein binding. Binding of tryptophane-containing peptides and peptides with free N-terminal amino acids can be eliminated by acetylation of the N-terminal amino acids [25].

The tryptic sample was applied onto a copper (Cu^{2+}) loaded IMAC column in order to select histidine-containing peptides, and thereby simplify the sample for further analysis with RPC, CE, and MALDI–TOF; the resulting IMAC chromatogram is shown in Fig. 4. Early eluted peptides are those with low affinity to the immobilised metal. In a previous study it was shown that these peptides were eliminated from the IMAC chromatogram, when N-terminal amino acids were acetylated [5]. MALDI analysis of the later eluted peaks showed that IMAC not only selected the histidine-containing peptides but also other peptides, perhaps on the basis of the free α-amino group, because cystein residues were alkylated. IMAC selected all the histidine-containing peptides, which were present in the tryptic sample (Table 2).

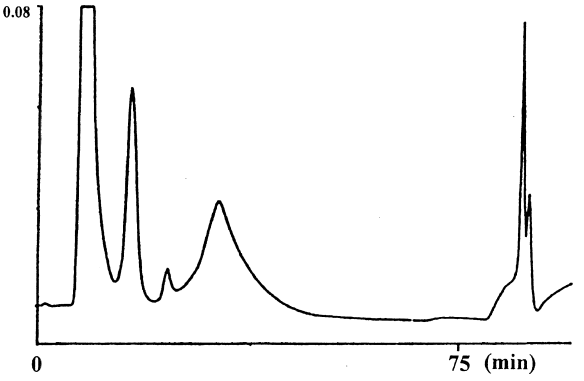

Fig. 4. IMAC analysis of the tryptic digest of ovalbumin. Conditions: See the Experimental section.

3.3. Analysis of IMAC fraction by capillary electrophoresis

To investigate the role of capillary electrophoresis off-line coupled to MALDI as one of the approaches to develop new strategies for separation and identification of signature peptides, the IMAC sample was analysed by CE under various conditions: (i) 25 m*M* phosphate buffer, pH 3.0 (Fig. 5a), (ii) 20 m*M* phosphate buffer, pH 9.0 (Fig. 5b), and (iii) 10 m*M* SDS dissolved in 25 m*M* ammonium acetate buffer, pH 7.0 (Fig. 5c). No significant change in the migration time of the peak was observed upon the addition of SDS into the BGE (compare Fig. 5b and 5c). This may depend on the low affinity of the analytes towards the SDS-micelles as a pseudo stationary phase and/or adsorption of SDS monomers to the capillary wall, which results in higher electroosmotic flow and shorter migration time. The peak obtained in the presence of 10 m*M* SDS (Fig. 5c) was successfully analysed by MALDI. The mass spectrum shows that the sample recovered from the capillary contains several peptides, among them one histidine-containing peptide. As Table 2 shows, the IMAC sample contained another histidine-containing peptide (HIATNAVLFFGR, M_w=1345.737), which could not be detected by CE, because of its low concentration. The molecular mass of other peptides does not match the data in Table 1. Possible explanations include the presence of impurities in the ovalbumin sample, miscleavages and also fragmenta-

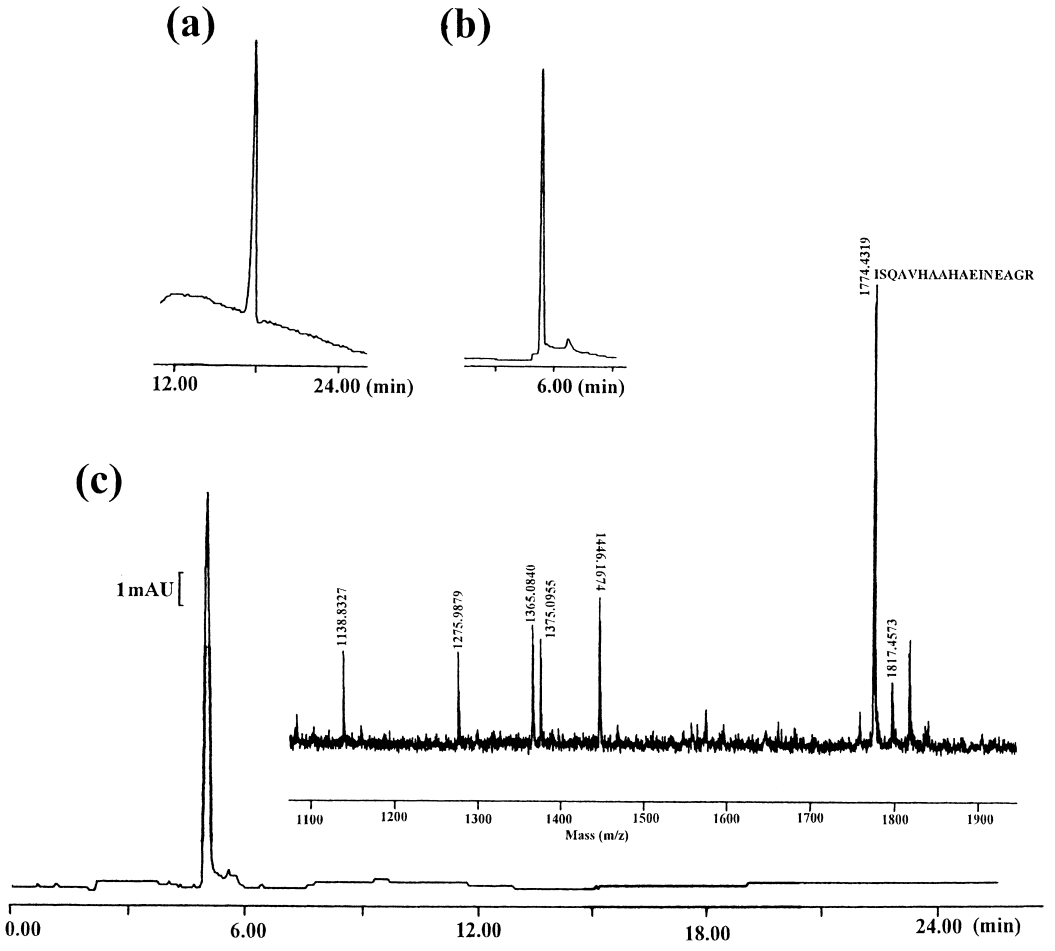

Fig. 5. CZE and MEKC analysis of the IMAC fraction. Operating conditions: BGE was (a) 25 m*M* phosphate buffer (pH 3.0); applied voltage was +15 kV, (b) 20 m*M* phosphate buffer (pH 9.0), and (c) 10 m*M* SDS dissolved 25 m*M* ammonium acetate buffer (pH 7.0); applied voltage was +10 kV. The MALDI spectrum was obtained from off-line analysis of the peak in electropherogram (c). Other conditions as those given in Fig. 1.

tion and loss of a water molecule or a hydroxyl group, during the MALDI analysis.

To separate the histidine-containing peptide from the matrix containing imidazole and other peptides, the concentration of SDS was increased to 75 m*M* and pH of the BGE was adjusted to 9.0, in order to eliminate electrostatic interaction between the imidazole molecules and the pseudo stationary phase. The analysis of the sample produced three peaks (Fig. 6b). MALDI analysis of the peaks at such a high SDS concentration is impossible [14]. Therefore, for identification of the peaks, the IMAC sample was run on a reversed-phase column. Two major fractions were collected for further analysis with MEKC. The MALDI analysis of the RP-fractions is shown in Fig. 7. Fraction 1 contained four peptides, whose mass could not be matched with any of the peptides listed in Table 1. The second fraction was found to be a mixture of the histidine-containing peptide and an unknown peptide at low concentration. Electropherograms obtained from MEKC analysis of the RP-fractions are shown in Fig. 6c and 6d. Each RP-fraction was resolved into two peaks by MEKC. The histidine-containing peptide and the other pep-

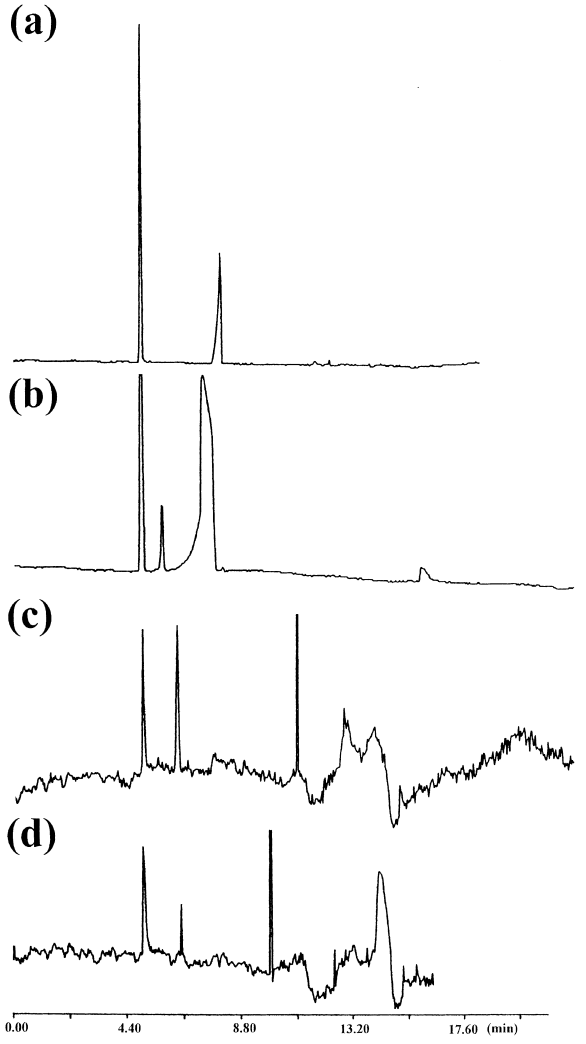

Fig. 6. Analysis of IMAC and RP-HPLC fractions by capillary electrophoresis with MEKC. Operating conditions: BGE was 20 m*M* phosphate buffer (pH 9.0) containing 75 m*M* SDS; applied voltage was +10 kV; other conditions as in Fig. 1. (a) IMAC eluting buffer containing 20 m*M* imidazole and 500 m*M* NaCl; (b) IMAC sample; (c) and (d) fractions 1 and 2 from RPC analysis of the IMAC sample.

tide present in fraction 2 co-migrated with imidazole. The results indicate that distribution of the peptides into the micelles is low, which may depend on the hydrophilicity and the charge of the peptides, being negatively charged at pH 9.0. The relatively poor limit of detection (LOD) achievable by CE makes it impossible to detect all peptides present in the IMAC and the RPC samples, which makes it difficult to evaluate the separation efficiency of MEKC.

It must be mentioned that tryptic digests of α_1-acid glycoprotein, α-amylase, pepsin and maltase were also analyzed by IMAC and CE. The CZE-analysis of the IMAC fractions resulted in a single peak, containing several peptides, analyzing the peaks by MALDI (data not shown).

4. Conclusions

IMAC was shown to be an efficient chromatographic technique to select histidine-containing peptides and simplify the complex tryptic samples. Two of five expected histidine-containing peptides from a tryptic digest ovalbumin were recovered. One of the peptides was identified analysing the IMAC sample and the RP fractions of the IMAC sample by capillary zone electrophoresis and micellar electrokinetic chromatography, respectively. Compared with CZE, MEKC proved to be a better technique for separation of the IMAC and reversed-phase fractions.

A fraction collected from IMAC or reversed-phase consists of a mixture of peptides. The electrophoretic mobility of the selected peptides seems to be very close, indicating that the charges on the peptides compensate for the mass and size variations resulting in the same charge-to-mass ratio. However, some of the peptides were separated employing MEKC at high micelle concentration; the drawback is that MEKC with high surfactant concentration (>20 m*M*) cannot be coupled to MALDI–TOF.

Acknowledgements

This research was support by grants from the National Institute of Health (GM 59996) and PE-Biosystems. A.A. gratefully acknowledges support from the Swedish Academy of Pharmaceutical Sciences (The Göran Schill Memorial Foundation) and the I.F. Foundation of Pharmaceutical Research.

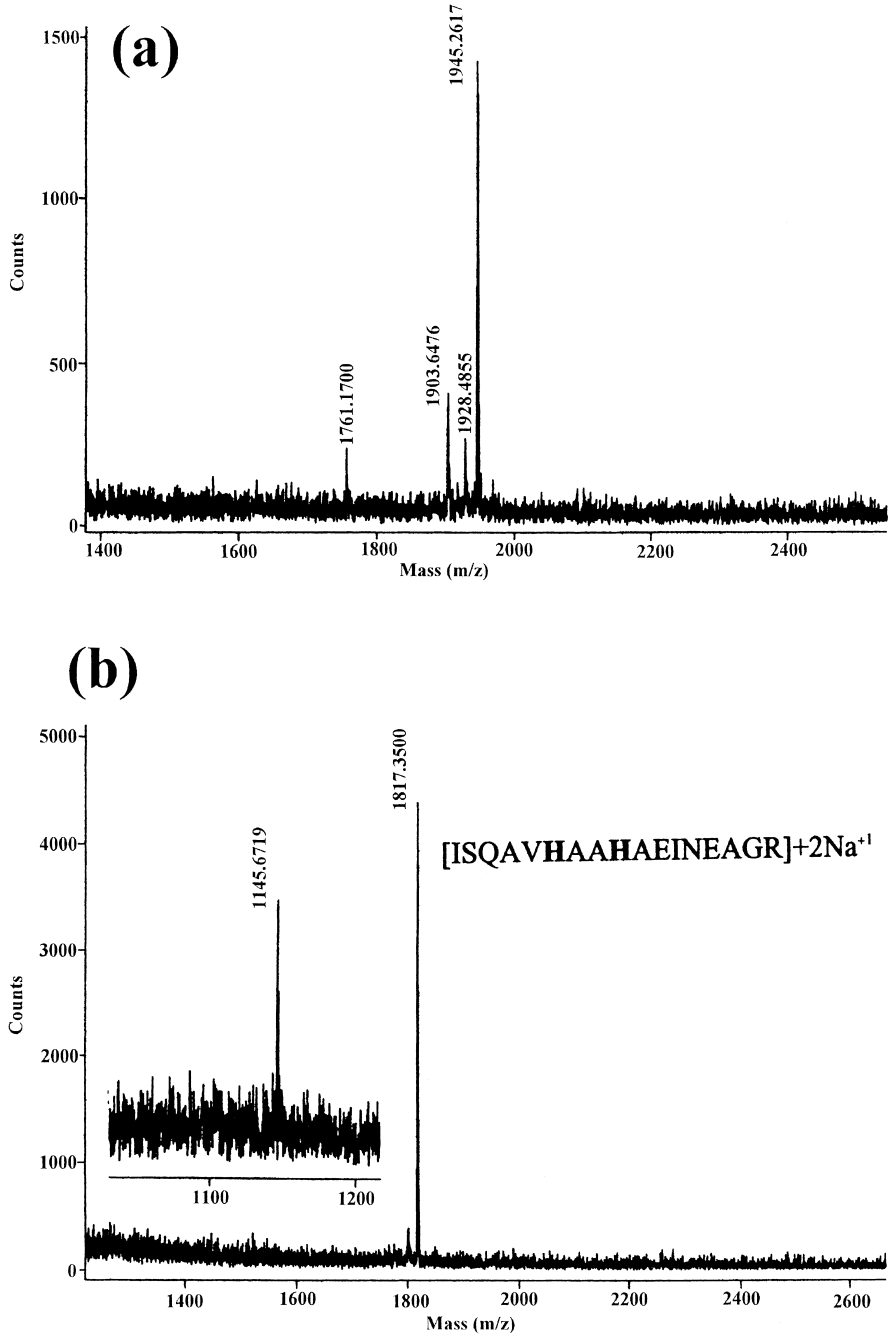

Fig. 7. MALDI analysis of RPC fractions. Conditions as in Fig. 3. (a) Fraction 1 and (b) fraction 2.

References

[1] M.J. Wise, T.G. Littlejohn, I. Humphery-Smith, Electrophoresis 18 (1997) 1399.
[2] J. Porath, J. Carlsson, I. Olsson, G. Belfrage, Nature 258 (1975) 598.
[3] E.S. Hedman, J. Porath, J. Chromatogr. 318 (1985) 265.
[4] E.S. Hedman, J. Porath, J. Chromatogr. 323 (1985) 247.
[5] J. Ji, A. Chakraborty, M. Geng, X. Zhang, A. Amini, M. Bina, F. Regnier, J. Chromatogr. B 745 (2000) 197.
[6] S. Terabe, K. Otsuka, K. Ichikawa, A. Tsuchiya, T. Ando, Anal. Chem. 56 (1984) 111.
[7] J. Lie, K.A. Cobb, M. Novotny, J. Chromatogr. A 516 (1990) 189.
[8] N. Matsubara, S. Terabe, Chromatographia 34 (1992) 493.
[9] U. Yashima, A. Tsuchiya, O. Morita, Anal. Chem. 64 (1988) 2299.
[10] M.G. Khaledi (Ed.), High Performance Capillary Electrophoresis, Wiley, New York, 1998, p. 77.
[11] M. Karas, D. Bachmann, U. Bahr, F. Hillenkamp, Int. J. Mass Spectrom. Ion Processes 78 (1978) 53.
[12] K. Tanaka, H. Waki, Y. Ido, S. Akita, Y. Youshida, T. Youshida, Rapid Commun. Mass Spectrom. 8 (1988) 151.
[13] V. Redecker, J.-Y. Toullec, J. Vinh, J. Rossier, D. Soyez, Anal. Chem. 70 (1998) 1805.
[14] A. Amini, S.J. Dormady, L. Riggs, F.E. Regnier, J. Chromatogr. A 894 (2000) 345.
[15] A.D. Nisbet, R.H. Saundry, A.J.G. Moir, L.A. Fothergill, J.E. Fothergill, Eur. J. Biochem. 115 (1981) 335.
[16] J.A. Chakel, E. Pungro Jr., W.S. Hancock, S.A. Swedberg, J. Chromatogr. B 689 (1997) 215.
[17] A. Gottschalk (Ed.), Glycoproteins: Their Composition, Structure and Function, Elsevier, Amsterdam, 1972, p. 732.
[18] R.P. Oda, B.J. Maden, T.C. Spelsberg, J.P. Landers, J. Chromatogr. A 680 (1994) 85.
[19] N. Ramadan, J. Porath, J. Chromatogr. 321 (1985) 93.
[20] H. Sigel, R.B. Martin, Chem. Rev. 82 (1982) 385.
[21] T.-T. Yip, Y. Nakagawa, J. Porath, Anal. Biochem. 183 (1989) 159.
[22] P. Hansen, G. Lindberg, J. Chromatogr. A 662 (1994) 235.
[23] P. Hansen, L. Anderssn, G. Lindberg, J. Chromatogr. 732 (1996) 51.
[24] P. Hansen, G. Lindberg, L. Andersson, J. Chromatogr. A 627 (1992) 125.
[25] R.J. Todd, R.D. Johnson, F.H. Arnold, J. Chromatogr. A 662 (1994) 13.