ELSEVIER

Journal of Chromatography A, 672 (1994) 219–229

JOURNAL OF CHROMATOGRAPHY A

Comparison of the utility of capillary zone electrophoresis and high-performance liquid chromatography in peptide mapping and separation

Suzanne E. Rudnick[a], Vincent J. Hilser, Jr.[b], Gregory D. Worosila*[,c]

[a]*Department of Chemistry, Manhattan College, Riverdale, NY, USA*

[b]*Graduate Program in Biology, Johns Hopkins University, Baltimore, MD, USA*

[c]*Analytical Development Department, Ciba Pharmaceuticals Division, Ciba-Geigy Corporation, Suffern, NY 10901, USA*

(First received October 18th, 1993; revised manuscript received March 2nd, 1994)

Abstract

Capillary zone electrophoresis (CZE) and high-performance liquid chromatography (HPLC) have been used in the analysis of the primary structure of recombinant human insulin-like growth factor I (rhIGF-I). CZE both complements and supplements HPLC separations. CZE has been used to resolve peaks which co-elute on HPLC, as well as to help establish the identity of tryptic fragments in peptide mapping experiments.

1. Introduction

Capillary zone electrophoresis (CZE) has become a major new analytical tool in chemistry and biochemistry laboratories. CZE separations are performed in free solution with no solid-phase matrix, facilitated by use of a capillary tube (100 cm × 100 μm) and conditions of high voltage (30 kV) and low current (50 μA). CZE has allowed separations previously unattainable and improved separations in already existing systems [1]. In addition, certain characteristics intrinsic only to CZE separations allow this technique to be used in situations where other separation techniques are not applicable [2].

We report here that CZE can both complement and supplement HPLC, particularly in peptide mapping experiments. Specifically, we have performed peptide mapping experiments on recombinant human insulin-like growth factor I (rhIGF-I), using both HPLC and CZE. Consequently, we have compared HPLC analysis of tryptic fragments of rhIGF-I with CZE analysis of the same peptides. This comparison has helped to elucidate the relative applicability of these two techniques.

As a complementary method, CZE is useful in ensuring the integrity of peptide fragments collected from HPLC analysis of enzymatic digests. Reversed-phase (RP) HPLC separation of proteins and peptides is generally achieved via interactions between specific non-polar amino acid residues and the column stationary phase, so that factors other than the side-chain residues may affect the separation [3]. CZE separations depend on peptide solvent interactions [4], where physical characteristics such as charge-to-

* Corresponding author.

0021-9673/94/$07.00
SSDI 0021-9673(94)00212-R

mass ratios are important [5]. CZE may thus separate closely related peptides which may coelute in the RP-HPLC mode. As a supplemental method, we demonstrate here the use of CZE for the positive identification of peptide peaks on an electropherogram.

2. Experimental

rhIGF-I was obtained from Ciba Pharmaceuticals. A single batch of rhIGF-I, which was >98% chromatographically pure by protein mass, was used to generate peptide fragments. HPLC-grade acetonitrile (Fisher, Pittsburgh, PA, USA) was used for HPLC analysis of tryptic fragments, optima-grade acetonitrile (Fisher) was utilized for all amino acid analysis experiments, and certified trifluoroacetic acid (TFA) (Fisher) was used for all HPLC analyses. In addition, reagent grade iodoacetamide (Sigma, St. Louis, MO, USA), gold-label triethylamine (Aldrich, Milwaukee, WI, USA), and sequanal-grade phenylisothiocyanate (PITC) and constant-boiling HCl (Pierce, Rockford, IL, USA) were used. All other chemicals were analytical grade and not further purified. Distilled, deionized Milli-Q water (Millipore, Bedford, MA, USA) was used for all experiments. Trypsin, in various treated forms, was purchased from three sources (Sigma, Pierce and ICN, Cleveland, OH, USA) and used without further purification.

2.1. Microdialysis

A Pierce Model 500 micro dialyzer with a molecular mass cutoff of 1000 was used.

2.2. RP-HPLC apparatus

Chromatographic separations were performed with a system composed of a Waters (Milford, MA, USA) Model 600E gradient pump and WISP Model 712 autosampler with temperature control, a Kratos (Foster City, CA, USA) Spectroflow 783 UV detector at 214 nm, and a Dionex (Sunnyvale, CA, USA) eluent degas module. Data collection and peak processing were performed with a Nelson 760 Series interface and a modified version of Nelson analytical software. A Gilson (Model 201) programmable fraction collector was used to isolate HPLC fractions. Well resolved peaks were collected by time programming, while closely eluting peaks were collected manually to avoid contamination due to small shifts in migration times.

2.3. Capillary electrophoresis apparatus

CZE was performed on an Applied Biosystems 270A capillary electrophoresis system with a fused-silica capillary (122 cm × 100 μm). Ultraviolet detection at 200 nm was used for peak analysis. Data collection and peak processing were performed as described for the HPLC analysis.

2.4. Amino acid analysis

Peptides were identified by amino acid analysis. The Waters Pico-Tag work station was utilized for the gas-phase hydrolysis of each peptide fragment and subsequent generation of PITC-derivatized amino acids [6]. Separation of derivatized amino acids, data collection and peak processing were performed by HPLC, as described below.

2.5. Carboxymethylation

Iodoacetamide was used for the carboxymethylation of rhIGF-I according to standard methods [7]. After derivatization, microdialysis was used for buffer exchange.

2.6. Trypsin digestion of rhIGF-I

The hydrolysis conditions utilized were a modification of those described by Worosila [8]. A 10-μl volume of trypsin solution (10 mg/ml in 0.1 m*M* HCl) was added to a solution of carboxymethylated rhIGF-I (250 μl of a 10 mg/ml solution in water) initially and, again, after 3 h (trypsin–rhIGF-I 1:25). After an 18-h incubation at 37°C, the reaction was quenched by the addition of 100 μl of 10% (v/v) TFA in water.

2.7. RP-HPLC method

Reversed-phase chromatographic analyses of rhIGF-I tryptic digests were performed on a Vydac Protein and Peptide C_{18} column (15.0 cm × 4.6 mm I.D.) (Vydac, Hesperia, CA, USA) at ambient temperature. The flow-rate was 0.8 ml/min. Mobile phases consisted of (A) 0.1% TFA in water and (B) 0.08% TFA in acetonitrile–water (80:20). A gradient was employed which ran according to the following program: 100% A at 3 min, 65% A at 38 min, 0% A at 55 min. Final gradient conditions were kept for 5 min before being returned to the initial conditions. A 15-min equilibration time was utilized between each run. The injection volume was 20 μl for optimum resolution and 30 μl for fraction collection.

2.8. Peak identification

Peptides were identified using amino acid analysis by comparison with a standard mixture. HPLC fractions were pooled from ten runs and lyophilized in acid-washed vials. Peak purity was determined by rechromatographing collected peaks using the gradient elution described above. The lyophilisate was then redissolved in 200 μl of distilled water and 30 μl used for amino acid analysis using a modification of the Waters Pico-Tag PITC method. Samples were dried under vacuum, subjected to gas-phase hydrolysis for 18 h at 105°C, and again dried under vacuum and made alkaline with a redrying solution [ethanol–triethylamine–water (40:20:40, v/v/v)]. Redrying was repeated and the samples were combined with 20 μl of derivatizing solution [ethanol–PITC–triethylamine–water (70:10:10:10, v/v)]. Solutions were allowed to react for 20 min at ambient temperature, again dried under vacuum, diluted in 200 μl of sample diluent [0.05 *M* disodium hydrogenphosphate, pH 7.4 in acetonitrile–water (5:95, v/v)] and analyzed chromatographically. RP-HPLC analyses of amino acid derivatives were performed using a Waters Pico-Tag amino acid analysis column maintained at 38°C. The flow-rate was kept at 1.0 ml/min for 12 min and changed linearly to 1.5 ml/min over a 30-s period. Mobile phases consisted of (A) 0.13 *M* sodium acetate trihydrate in triethylamine, pH 6.4–acetonitrile–water (0.05:6:93.95, v/v/v) and (B) acetonitrile–water (60:40, v/v). A concave gradient was employed which ran according to the following program: 100% A to 54% A at 10.0 min, 0% A at 10.5 min.. Final mobile phase conditions were kept for 1 min and the flow gradient listed above was overlaid on the mobile phase gradient. The column was equilibrated at a flow of 1.5 ml/min for 8 min between injections. The injection volume was varied to optimize peak response and the amino acid derivatives were detected at 254 nm.

2.9. CZE method

CZE was performed on all peptide peaks isolated from RP-HPLC. After lyophilization, the samples were dissolved in 10 m*M* sodium citrate, pH 2.5 at a concentration of approximately 0.5 mg/ml. The CZE mobile phase was 20 m*M* sodium citrate, pH 2.5. CZE was performed at 30°C with a constant voltage of 30 kV. Samples were introduced by a 1-s vacuum injection onto a 122-cm capillary (99.7 cm to the detector). A proprietary mobility standard was utilized to calculate peptide mobilities, thereby eliminating drift in the elution times. Electrophoretic mobilities (μ) of samples were calculated according to Eq. 1.

$$\mu = \frac{(L_d L_t)}{V} \cdot \left(\frac{1}{t} - \frac{1}{t_s}\right) + \mu_s \quad (1)$$

where L_d = length of capillary to detector (cm), L_t = total length of capillary (cm), V = system voltage (V), t = migration time of sample peak (s), t_s = migration time of standard peak (s) and μ_s = mobility value for standard peak.

2.10. Charge calculation

All peptide charges were calculated using the Henderson–Hasselbach equation using the pK_a values in ref. 9.

3. Results

3.1. Peptide fragment identification

The primary sequence of rhIGF-I is shown in Fig. 1 along with the nine theoretical tryptic cleavage sites. The ten resulting tryptic fragments are labelled T1–T10. Table 1 shows the amino acid molar ratios of all of the peptide fragments isolated by HPLC (see Fig. 3), which were used to identify the corresponding peptide stretches. Residual chymotryptic activity, which is often seen in tryptic digests, was evident by the appearance of some non-tryptic fragments. These fragments were used in the calculations described in this paper. Since the primary structure of the current batch of rhIGF-I had previously been determined by fast-atom bombardment mass spectrometry, the amino acid ratios were sufficient to identify the peaks [10].

In addition to amino acid analysis, aliquots of all fractions were analyzed by CZE. Comparison of the retention behavior in the two methods formed the basis for peak assignment in all comparative chromatograms. In all but one case, fractions which eluted as one peak on HPLC also proved to be homogenous by CZE analysis. In the case of coeluting peaks on RP-HPLC, positive identification was accomplished through the use of amino acid analysis and electrophoretic mobility on CZE.

3.2. Method development

Trypsin digestions are routinely performed in urea-containing solutions to minimize precipitation; however, the desire to analyze the resultant fragments by CZE limited the amount and type of buffer used. Although buffer com-

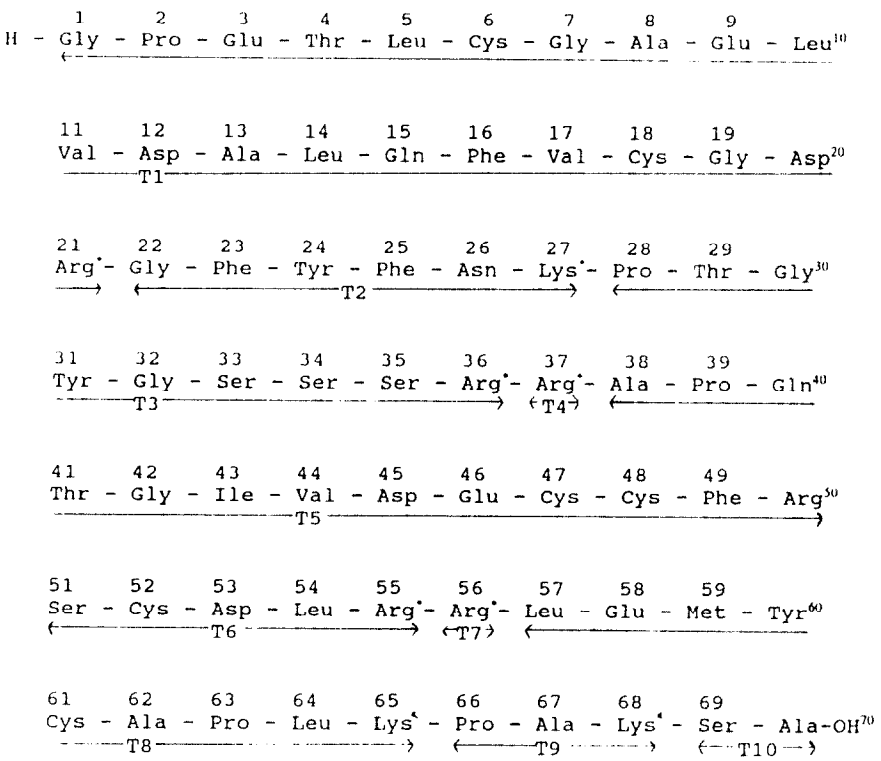

Fig. 1. The nine theoretical tryptic cleavage sites of rhIGF-I. Arrows in combination with an asterisk represent cleavage sites at the carboxyl terminal of Arg and Lys residues. Tryptic fragments are labeled T1 through T10.

Table 1
The amino acid analysis of rhIGF-I

Amino acid	Molar ratio														
	Gly1–Leu14	Gly1–Arg21	Gln15–Phe16	Val17–Arg21	Gly22–Tyr24	Gly22–Phe25	Phe25–Tyr31	Phe25–Arg36	Arg37–Arg50	Ser51–Arg55	Arg56–Tyr60	Arg56–Ala70	Leu57–Tyr60	Cys61–Lys65	Pro66–Lys68
Asx	1.0	1.8		1.0			0.8	1.2	1.0	1.1					
Glx	2.2	3.3	1.0						2.0		1.0	1.0	1.3		
CM-Cys	P[a]	P		P			P		P	P				P	
Ser								2.7		1.0		0.3			
Gly	2.3	3.2		0.9	1.1	1.3	1.2	1.9	1.2						
Arg		0.9		1.0			1.3	1.8	1.0	1.2	1.3				
Thr	1.0	1.0					1.1		1.0						
Ala	1.9	1.8						0.8	0.9			2.4		1.0	1.0
Pro	1.1	1.0					1.4	0.6	1.0			2.4		1.2	1.0
Tyr					0.7	1.0	0.9				0.9	0.6	0.9		
Val	1.0	2.5		0.8					0.6		0.8	0.5	0.9		
Met															
Ile									0.6						
Leu	2.8	3.4								0.9	0.9	1.9	1.0	0.5	
Phe		1.1	0.8		1.0	1.8	1.1	1.0	0.9						
Lys							1.4	0.8				2.3		1.2	1.2
Fragments		T1			T2A			T2BT3	T4T5	T6		T8			T9

The amino acid analysis was performed as described in ref. 5. PITC-labelled amino acids were quantified by comparison with a standard mixture of the 20 common amino acids, after separation by HPLC.
[a] P = peak present, but not quantified.

ponents elute early and consequently do not interfere with sample peaks in HPLC, many buffer ions exhibit mobilities similar to samples in CZE and must therefore be used in minimal quantities. In order to increase the solubility of intermediate-sized fragments, and at the same time eliminate buffer additives, the cysteine residues were carboxymethylated with iodoacetamide. The ability to irreversibly reduce disulfide bonds and dialyze reagents out of the sample buffer allows for minimal baseline noise in CZE.

A standard for the carboxymethylated cysteine was not prepared, and the location of the PITC carboxymethyl-cysteine was determined by comparison of the total amino acid hydrolysates for the reacted and unreacted protein (Fig. 2). The extent of the carboxymethylation reaction was determined by amino acid analysis through the concomitant appearance of a peak thought to be the carboxymethylated cysteine with the disappearance of peaks corresponding to unreacted cysteine and cystine residues (Fig. 2). The predicted trypsin cleavage sites within the amino acid sequence of rhIGF-I are indicated in Fig. 1. Of the nine potential cleavage sites, six are arginine and three are lysine residues. Analysis of the sequence reveals two points of arginine tandem repeats (Arg-36,37 and Arg-55,56). The poor exoproteolytic ability of trypsin results in a mix of peptide fragments in these regions.

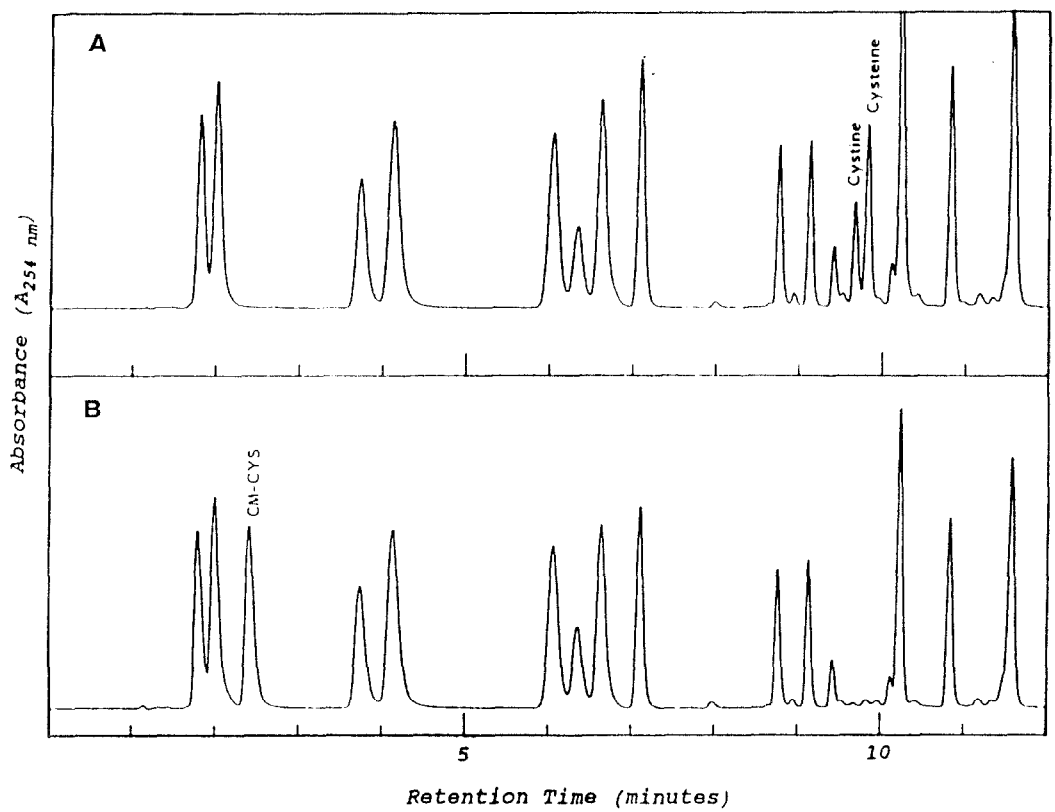

Fig. 2. Comparison of amino acid analysis results of unreacted (A) and derivatized rhIGF-I (B). The figure shows HPLC chromatograms of component amino acids. The disappearance of the cysteine and cystine peaks in (B) is concomitant with the appearance of the peak corresponding to the carboxymethylated cysteine (CM-CYS). Absorbance range in both chromatograms is 65 mAU.

3.3. Analysis of tryptic map of rhIGF-I

Tryptic hydrolysis yielded the HPLC map shown in Fig. 3. Amino acid analysis results indicated that the peaks corresponding to fragments T8, T7T8, T7T8T9 and T8T9T10 coelute on the HPLC system at 34 min, and result from the tandem arginine repeats. The coelution of the four fragments indicates that retention on the reversed-phase medium of the HPLC column is heavily influenced by the hydrophobic stretches of the T8 fragment since this sequence is common to all four fragments. CZE analysis of the HPLC fraction collected at 34 min showed four major peaks of the following area percentages: peak 1, 59%; peak 2, 22%; peak 3, 11%; peak 4, 6% (Fig. 4A). This observation was investigated in an attempt to further delineate the differences in HPLC and CZE as laboratory techniques for peptide separation. Degradation and cross-contamination were not the cause of the impurity profile demonstrated on CZE as the fraction was subsequently reanalyzed by HPLC and found to be >98% chromatographically pure (Fig. 4B).

3.4. Selectivity of CZE and HPLC separations

That the separation mechanisms of CZE and HPLC are different is seen by the behavior of the T8-containing fragments (T8, T7T8, T7T8T9 and T8T9T10). HPLC separations involve a mass transfer which is energetically determined by the hydrophobic character of the sample. Since most

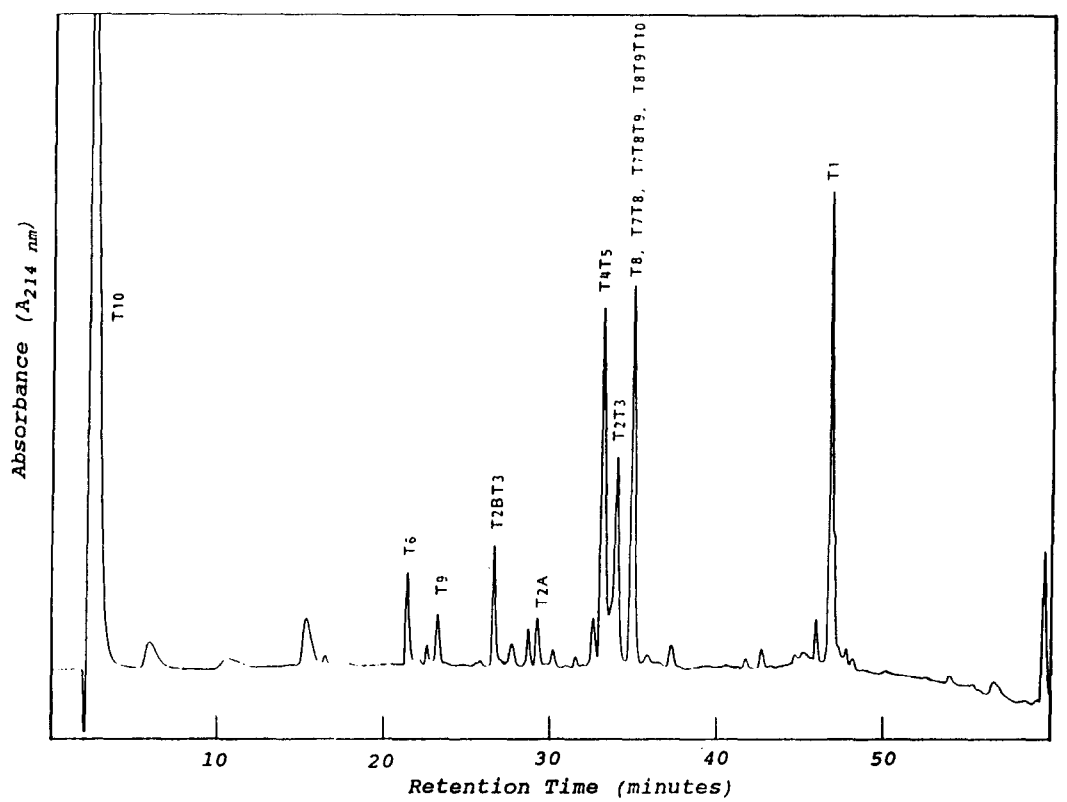

Fig. 3. Final HPLC tryptic map of rhIGF-I. Incomplete digestion products are indicated by continuous fragment labels (*e.g.* T3T4). Multiple coeluting peaks are separated by commas. A and B are used to indicate two parts of a tryptic fragment which were cleaved at a non-tryptic cleavage site (*e.g.* T2A and T2BT3). The absorbance range is 500 mAU.

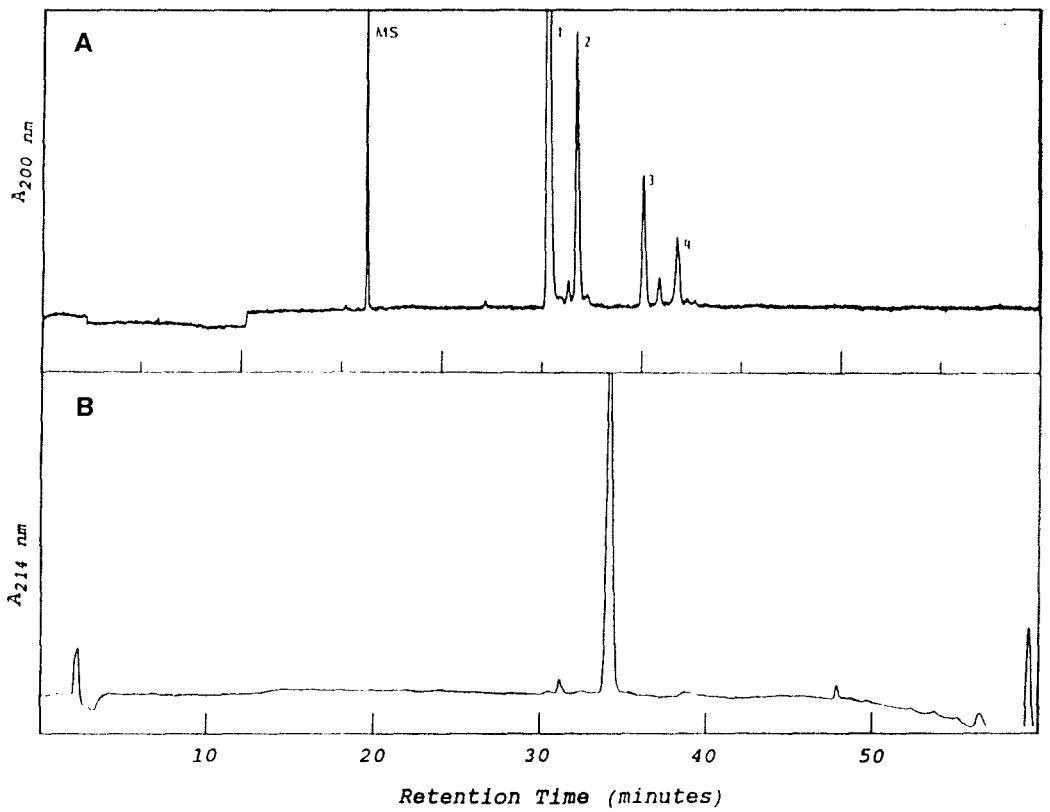

Fig. 4. Comparison of the CZE (A) and HPLC (B) analysis of the T8-containing fragments. The identities of these fragments are: peak 1 = T7T8; peak 2 = T7T8T9; peak 3 = T8T9T10; peak 4 = T8. MS = Mobility standard. The absorbance range is 7 mAU for CZE and 500 mAU for HPLC.

peptides are intrinsically amphipathic, separation is most often a function of a limited region of a peptide. CZE separations depend on the sample charge and size. Because mass transfer does not occur in CZE, mobility should be influenced by each residue in the peptide. If all chemical moieties contribute to mobility, it should be possible to separate very similar peptides. Further resolution by CZE of the fraction containing the T8 fragments illustrates this point.

A comparison of the tryptic map of rhIGF-I was performed to elucidate the differences in selectivity between CZE and HPLC. Fig. 5 indicates a large difference in the selectivity of both methods, where lines are used to connect identical peaks and to demonstrate reversals in the elution order between the two separation mechanisms.

Amino acid analysis of reversed-phase separated T8-containing fragments showed molar ratios suggestive of the expected fragments; however, two anomalies in the amino acid analysis rendered the ratios inconclusive —the molar ratios calculated for alanine, proline, lysine and arginine were not stoichiometric (Table 1), and a low recovery was calculated for serine (0.3 molar ratio). Through a systematic analysis of possible peptide cleavages, it was determined that the CZE peaks in Fig. 4A should have the following amino acid sequences. However, these data alone were insufficient to predict the relative elution order of these fragments.

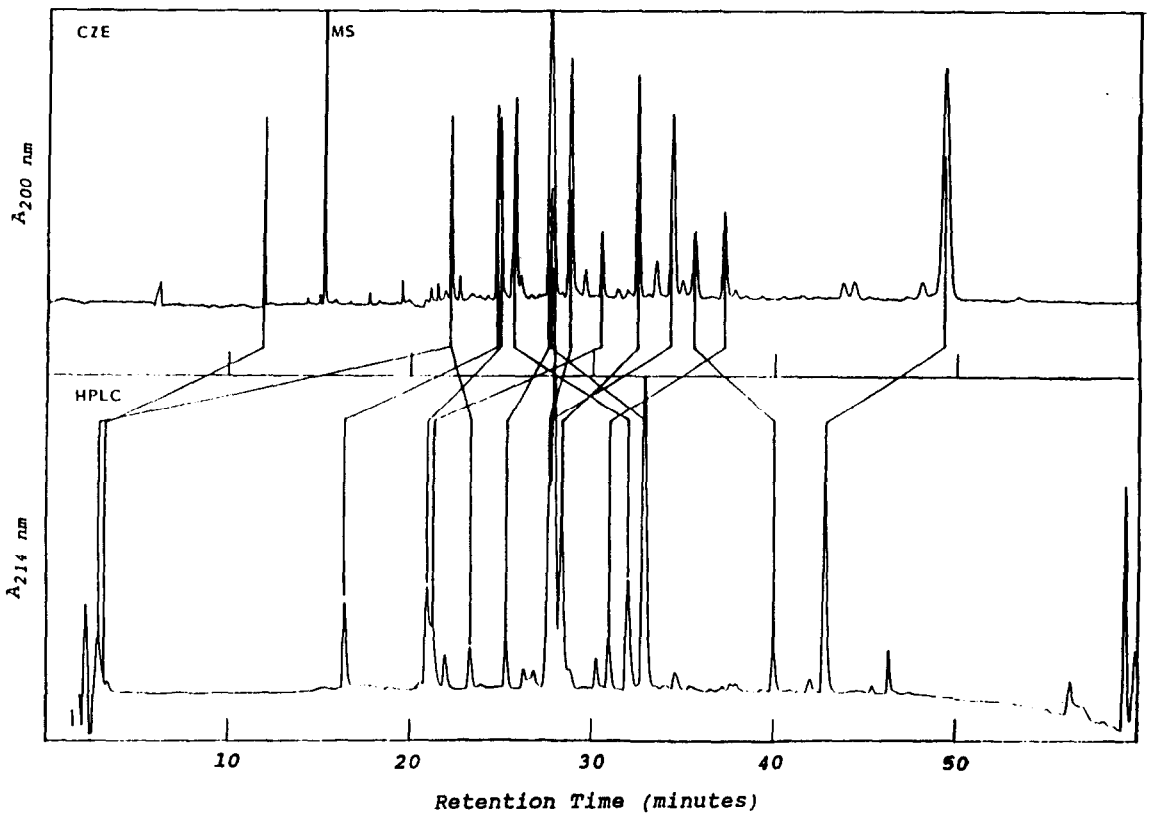

Fig. 5. Comparison of selectivity between CZE and HPLC. Lines are used to connect identical peaks and emphasize changes in elution order. MS = Mobility standard. The absorbance range is 7 mAU for CZE and 500 mAU for HPLC.

Peak	*T7*	*T8*	*T9*	*T10*
1	Arg–	Leu–Glu–Met–Tyr–Cys–Ala–Pro–Leu–Lys		
2	Arg–	Leu–Glu–Met–Tyr–Cys–Ala–Pro–Leu–Lys–	Pro–Ala–Lys	
3		Leu–Glu–Met–Tyr–Cys–Ala–Pro–Leu–Lys–	Pro–Ala–Lys–	Ser–Ala
4		Leu–Glu–Met–Tyr–Cys–Ala–Pro–Leu–Lys		

Since the T8-containing fragments could not be unequivocally identified through routine amino acid analysis, it was decided to calculate the CZE electrophoretic mobilities using the polymer function (Eq. 2) proposed by Grossman *et al.* [11].

$$\mu \propto \ln \frac{(q+1)}{N^{0.43}} \qquad (2)$$

Table 2 shows the values used in Eq. 2, as well as experimental mobilities for each of the peptide fragments. In order to identify the elution order of the four T8-containing fragments, the polymer function in Eq. 2 was plotted against the experimental mobility values for each of the identified tryptic fragments (Fig. 6). The slope and intercept values from this plot were determined and then used to calculate the theoretical mobility value for each of the T8-containing fragments. The theoretical mobilities were then compared to the experimentally determined mo-

Table 2
Data for tryptic fragments of rhIGF-I

Peptide fragment	Charge (q)	Polymer number (N)	Molecular mass	Mobility (μ) ($\times 10^4$)
Leu57–Tyr60	0.31	4	554.7	1.31
Gly22–Tyr24	0.33	3	385.4	1.58
Gly1–Leu14	0.31	14	1443.6	0.90
Gln15–Phe16	0.43	2	293.3	1.74
Gly22–Phe25	0.43	4	532.6	1.41
Gly1–Arg21	1.06	21	2305.6	1.18
Val17–Arg21	1.16	5	604.6	2.25
Ser51–Arg55	1.16	5	648.7	2.23
Leu57–Lys65*	1.30	9	1123.4	1.91
Arg56–Tyr60	1.31	5	710.9	2.14
Pro66–Lys68	1.32	3	314.4	2.60
Cys61–Lys65	1.32	5	586.7	1.86
Phe25–Tyr31	1.33	7	825.9	1.97
Arg37–Arg50	2.13	14	1706.9	1.94
Arg56–Lys65*	2.31	10	1279.9	2.31
Leu57–Ala70*	2.39	14	1577.9	2.03
Phe25–Arg36	2.20	12	1300.4	2.21
Arg56–Lys68*	3.30	13	1576.0	2.46

Calculations of charge, polymer number, molecular mass and electrophoretic mobility for each of the peptide fragments obtained after tryptic digestion of rhIGF-I are summarized in the table. These data were subsequently used to obtain the plots in Figs. 6–8. Asterisks refer to T8-containing fragments.

bilities to establish the fragment identities. Once the peaks were identified, the polymer functions for the peptides were plotted against the actual mobility values (Fig. 7). The slope of the line derived from the four peptides correlated within 4% of that plotted for the identified tryptic fragments (Fig. 6). An overlay of Fig. 6 with Fig. 7 reveals a good correlation (Fig. 8) in which all

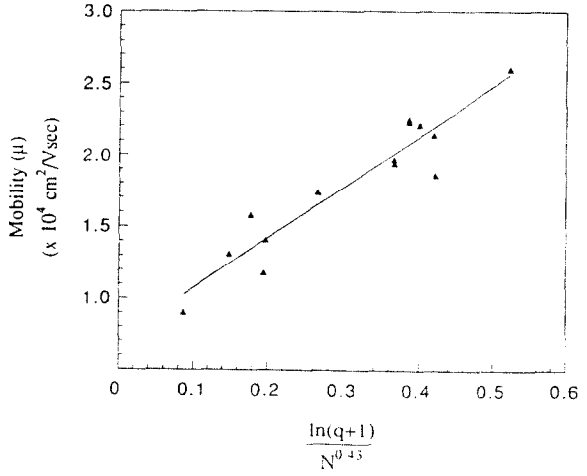

Fig. 6. Plot of the polymer function (Eq. 2) for all the peptides in Table 2 except the T8-containing fragment *versus* the mobility values.

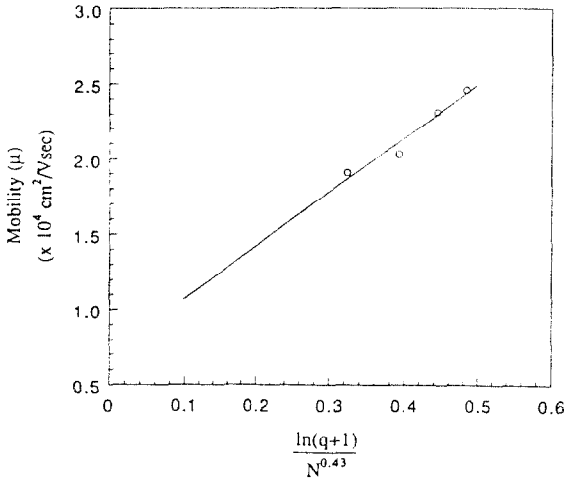

Fig. 7. Plot of the polymer function (Eq. 2) for the T8-containing fragments *versus* the experimental mobility values.

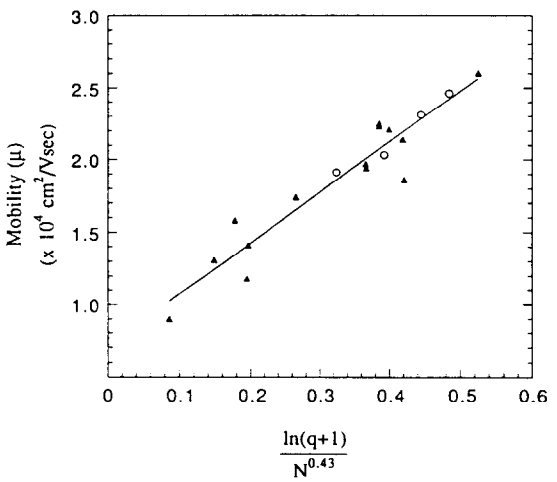

Fig. 8. Comparison of plots from Figs. 6 and 7.

four calculated points fall inside of the 95% confidence limits of the line. Validation of fragment assignment was derived by iteration of the assignments of the T8-containing fragments. The above peak assignments represent the only relationship in which all of the peaks fall within the 95% confidence limits of the line.

4. Discussion

Our analysis shows that CZE is a useful tool for both structural identification and purity confirmation in peptide mapping experiments. Although it is unlikely that CZE will take the place of HPLC for routine separation, its potential as a complementary method has been shown. The difference in the selectivity of CZE demonstrated has resulted in a peptide map which differs significantly from the HPLC map. In addition to the difference in selectivity, CZE offers the unique predictive ability associated with mobility correlations. The added potential, resulting from the mobility correlations, indicates a powerful supplemental role for CZE in peptide mapping experiments.

Acknowledgement

We thank Dr. Frank Gasparini for helpful discussions and critical assessment of this manuscript.

References

[1] R.G. Nielsen, R.M. Riggin and E.C. Rickard, *J. Chromatogr.*, 480 (1989) 393–401.
[2] P.D. Grossman, H.H. Lauer, S.E. Moring, D.E. Mead, M.F. Oldham, J.H. Nickel, J.R.P. Goudberg, A. Krever, D.H. Ransom, and J.C. Colburn, *Am. Biotechnol. Lab.*, 8 (1990) 35–43.
[3] B.L. Karger, L.R. Snyder and Cs. Horváth, *An Introduction to Separation Science*, Wiley, New York, 1973.
[4] J.W. Jorgensen and K.D. Lukacs, *Science*, 222 (1983) 266–272.
[5] V.J. Hilser, Jr., G.D. Worosila and S.E. Rudnick, *J. Chromatogr.*, 630 (1993) 329–336.
[6] B.A. Bidlingmeyer, S.A. Cohen and T.L. Tarvin, *J. Chromatogr.*, 336 (1984) 93.
[7] M.J. Waxdal, W.H. Konigsberg, W.L. Henley and G.M. Edelman, *Biochemistry*, 7 (1968) 1959–1966.
[8] G. Worosila, *Ph.D. Thesis*, Rutgers, State University of New Jersey, New Brunswick, NJ, 1985, p. 45.
[9] B. Skoog and A. Wichman, *Trends Anal. Chem.*, 5 (1986) 82.
[10] F. Raschdorf, R. Dahinden, W. Maerki and W.J. Richter, *Biomed. Environ. Mass Spectrom.*, 16 (1988) 3.
[11] P.D. Grossman, J.C. Colburn and H. Lauer, *Anal. Biochem.*, 179 (1989) 28.