ELSEVIER

Journal of Chromatography A, 848 (1999) 417–433

JOURNAL OF CHROMATOGRAPHY A

Peptide mobility and peptide mapping in capillary zone electrophoresis
Experimental determination and theoretical simulation

George M. Janini*, Climaco J. Metral, Haleem J. Issaq, Gary M. Muschik

SAIC Frederick, NCI-Frederick Cancer Research and Development Center, P.O. Box B, Frederick, MD 21702, USA

Received 21 December 1998; received in revised form 5 March 1999; accepted 15 March 1999

Abstract

The electrophoretic mobilities of 58 peptides that varied in size from 2 to 39 amino acids and varied in charge from 0.65 to 7.82 are presented. The measurements were conducted at 22°C using a 10% linear polyacrylamide-coated column and a 50 m*M* phosphate buffer at pH 2.5. Excellent separation of peptides and highly reliable peptide maps of protein digests are routinely obtained using these experimental conditions. The electrophoretic data were used to test existing theoretical models that correlate electrophoretic mobility with physical parameters. The results indicate that the Offord model that correlates electrophoretic mobility with the charge-to-size parameter $q/M^{2/3}$ offers the best fit of our reliable experimental data. Furthermore, we also obtained the capillary zone electrophoretic profile of the endoproteinase Lys-C digests of a peptide sequencing standard, melittin, and horse myoglobin under the same experimental conditions as described above. The resulting peptide maps were compared with corresponding theoretical simulation.

Keywords: Mathematical modelling; Peptides

1. Introduction

Stoke's law states that when a charged particle is placed in an electric field, it experiences a force that is proportional to its effective (net) charge (q) and the electric field strength (E). The translational movement of the particle is opposed by a viscous drag force that is proportional to the particle's velocity (v), hydrodynamic radius (r) and the viscosity of the medium (η). When the two forces are counterbalanced, the particle moves with a steady-state velocity [1]

$$\nu_{ef} = \mu_{ef}E = \mu_{ef}\frac{V}{L} \quad (1)$$

where V=applied voltage, L=column length and μ_{ef}=the electrophoretic mobility, which is given by:

$$\mu_{ef} = \frac{q}{4\pi\eta r} \quad (2)$$

Strictly speaking, Stoke's law is only valid for rigid, spherical molecules in low ionic strength buffers. The situation is not so simple for molecules that are large or even small, but not spherical. For example, for rod-shaped particles, the aspects ratio and the

*Corresponding author. Tel.: +1-301-846-1226; fax: +1-301-846-1438.

E-mail address: janini@ncifcrf.gov (G.M. Janini)

0021-9673/99/$ – see front matter
PII: S0021-9673(99)00388-X

orientation with respect to the direction of motion enter into play [2].

Several semi-empirical models have been advanced to explain the dependence of the electrophoretic mobility on charge (q) and size (r) [3–14]. Most models equate q with the valence of the molecule and relate r to the molar mass (M) or the number of amino acid residues (n). In summary, the various models suggest a direct dependence of μ_{ef} on q and a size dependency ranging from $1/r$ for small molecules to $1/r^2$ for large molecules. Assuming that the molar mass is proportional to the molar volume, and that a molecule can be represented by a sphere of volume $=(4/3)\pi r^3$, then r can be replaced by $M^{1/3}$ and r^2 by $M^{2/3}$. The conditions under which these assumptions are valid are discussed by Cross and Cao [13]. Offord [15] argued that an ion moving through a conducting medium would experience a retarding force that is proportional to its surface area. This implies that the ion's electrophoretic mobility would be proportional to $1/r^2$ ($1/M^{2/3}$), rather than $1/r$ ($1/M^{1/3}$), as stipulated by Stoke's model. Compton [6] combined the two models in one which predicts that the molar mass dependency of mobility is a continuous function of $1/M^{1/3}$ to $1/M^{2/3}$. Small molecules in low ionic strength buffer are more closely correlated with $1/M^{1/3}$ while large molecules in high ionic strength buffer are better correlated with $1/M^{2/3}$. Molecules of intermediate size in medium-strength buffers show dependence on $1/M^{1/2}$ [6]. Taking a different approach, Grossman et al. [4] considered the peptide as a classical linear polymer with n amino acid residues and arrived at an equation that correlates μ_{ef} with ln $[(q+1)/\eta^{0.43}]$. The logarithmic dependence of q was introduced to compensate for electrostatic charge suppression, which becomes increasingly significant for highly charged peptides [4,11]. A modification to the classical linear model of Grossman et al. [4], was reported by Cifuentes and Poppe [11], who retained the logarithmic dependence of mobility on charge but substituted M for n for the size dependence.

Successful correlations using any of the approaches described above require an accurate determination of charge. The fundamental equation that is invariably used for the calculation of the net charge on a peptide is the Henderson–Hasselbach equation [16]. The use of this equation requires an accurate knowledge of the ionization constants of the amino acid residues. Since these values are not accurately known, most researchers resort to the use of the pK values of free amino acids [17], assuming that the ionization of each amino acid is not affected by its nearest neighbors in the molecule, or use corrected pK values [5,11]. This approach for the determination of q was criticized by Cifuentes and Poppe [11] because it neglects the statistical effect that occurs when more than one proteolytic group is present. It also neglects the electrostatic and steric effects of neighboring residues [11]. Notwithstanding, it remains the only practical approach for the determination of charge for a large set of peptides. In one report, corrected pK values compared favorably with experimentally determined pK values for a set of six peptides [18].

In this work, we experimentally measured the electrophoretic mobility of 58 peptides that varied in size from 2 to 39 amino acids and varied in charge from 0.63 to 7.82. This set is, by far, the largest of any that has been measured previously in a single laboratory [3–14]. The measurements were conducted at 22°C using a polyacrylamide-coated capillary and a 50 mM phosphate buffer at pH 2.5. The results were used to test existing models in order to identify the model that best fit our experimental data. The salient features, and the strength and weakness of each model, will be presented and discussed. Furthermore, we also obtained the capillary zone electrophoresis (CZE) profile of the endoproteinase Lys-C digests of a Sigma peptide sequencing standard, melittin, and horse myoglobin under the same conditions as mentioned above. The resulting peptide maps will be compared with corresponding theoretical simulations. The long-range objective of this study is to develop a computer program that simulates ideal CZE peptide maps of protein digests based on their known amino acid sequence. Operation with polyacrylamide-coated columns at pH 2.5 ensures that electroosmotic flow and adsorption to the column wall are negligible. More importantly, all proteins and peptide fragments will have a net positive charge and, thus, will all migrate in the same direction towards the detector at the cathodic end of the column [19,20].

2. Experimental

2.1. Materials

The peptides used in this study are listed in Table 1, together with their amino acid sequences, molar masses and calculated charges, at pH 2.5. The net charge of a peptide at pH 2.5 was calculated as the algebraic sum of all charged residues and carboxyl- and amino-terminals. Each R, K, H and the amino-terminal contribute a charge +1. Each D, E and the carboxyl-terminal contribute a partial negative charge, calculated using the Henderson–Hasselbach relation [16]. The pK_a values used were 3.20 for the carboxyl-terminal, 3.50 for D-residues and 4.50 for E-residues [5,21]. It is to be noted that these numbers are only approximate because pK_a values of amino acid residues in peptides are affected by the nature of their nearest neighbours. Peptides 1–24, 28, 34, peptide sequencing standard, melittin and horse myoglobin were from Sigma (St. Louis, MO, USA). Sequencing grade endoproteinase Lys-C was obtained from Boehringer Mannheim (Indianapolis, IN, USA). The rest were custom-synthesized by commercial laboratories for Dr. J.A. Berzofsky (NCI, NIH, USA). The buffer components were purchased from Fisher Scientific (Pittsburgh, PA, USA). Only one buffer system was used throughout. The buffer was made up of 50 m*M* phosphoric acid that was adjusted to pH 2.5 with triethylamine (TEA). All reagents were used as received. The fused-silica columns were purchased from Polymicro Technologies (Phoenix, AZ, USA).

2.2. Apparatus and procedures

A Beckman CZE model P/ACE 5510 equipped with a UV detector, an automatic injector, a fluid-cooled column cartridge and a System Gold data station were used in this study. All runs were performed at 200 nm and 22°C. The buffers were prepared fresh daily, passed through 0.2 μm nylon filters, and degassed. Injections were made using the pressure mode at 0.5 p.s.i. (1 p.s.i. = 6894.76 Pa). The capillary inlet and outlet vials were replenished after every ten injections. The columns were coated with a dense layer of 10% polyacrylamide, as described in detail elsewhere [22]. The polyacrylamide-coated columns provided stability and migration time reproducibility throughout the experiments to within 1% RSD. This was established by periodically monitoring the electroosmotic mobility and measuring the migration time of a set of reference solutes. The electroosmotic mobility (μ_{eo}) of CZE columns is usually determined by measuring the migration time of a neutral marker such as mesityl oxide. The value of μ_{eo} for CE columns ranges from $0–10\cdot10^{-4}$ cm^2 V^{-1} s^{-1}, depending on column surface chemistry, pH and buffer additives [23]. For coated columns with almost total suppression of electroosmotic flow, the migration time of mesityl oxide is too long an, therefore, its direct measurement may not be practical. A procedure for the determination of μ_{eo} for coated columns is described in the next section.

2.3. Determination of μ_{eo} for coated columns

In this study, we determined μ_{eo} for our in-house coated columns by a procedure that we will illustrate with examples for a column of 37 cm total length (L_{tot}) and 30 cm injector-to-detector length (l). Each of the data points reported below is an average of three measurements. First, mesityl oxide was injected and eluted through the column by low-pressure (0.5 p.s.i.) and its migration time was recorded (267 s). A second injection of mesityl oxide was separated at a running voltage, V, of 12 kV for a period, t_m of 30 min. During this voltage separation step, the mesityl oxide zone moved a distance l_{mo}. After aborting the voltage separation, the mesityl oxide zone was eluted through the column by low pressure and the time of the pressure elution step was recorded (234 s). l_{mo} was then determined as follows:

$$l_{mo} = 30 - \left(\frac{234 \cdot 30}{267}\right) = 3.708 \text{ cm} \quad (3)$$

μ_{eo} was then obtained from the relation:

$$\mu_{eo} = \frac{l_{mo}.L_{tot}}{t_m.V} = \frac{3.708 \cdot 37}{12\,000 \cdot 30 \cdot 60} = 0.635 \cdot 10^{-5} \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1} \quad (4)$$

A repetition of the above procedure at a separation

Table 1
List of peptides used in this study

Peptide no.	Peptide sequence	Amino acid residues	Charge at pH 2.5	Molar mass
1	DD	2	0.65	248.2
2	FD	2	0.74	280.2
3	EE	2	0.81	276.3
4	GG	2	0.83	132.1
5	AA	2	0.83	160.2
6	PG	2	0.83	172.0
7	VV	2	0.83	216.3
8	FG	2	0.83	222.1
9	FA	2	0.83	236.2
10	LL	2	0.83	244.3
11	FV	2	0.83	264.2
12	FL	2	0.83	278.2
13	MM	2	0.83	280.4
14	FF	2	0.83	312.1
15	YY	2	0.83	344.4
16	WW	2	0.83	390.4
17	AAA	3	0.83	231.3
18	SSS	3	0.83	279.3
19	FFF	3	0.83	459.2
20	AAAA	4	0.83	302.3
21	AAAAA	5	0.83	373.4
22	YGGFL	5	0.83	555.5
23	YGGFM	5	0.83	573.5
24	RPPGF	5	1.83	572.6
25	AAGIGILTV	9	0.83	813.9
26	YMDGTMSQV	9	0.74	1030.4
27	VLQELNVTV	9	0.82	1014.2
28	RPPGFSPFR	9	2.82	1042.1
29	AFLPWHRLF	9	2.83	1186.4
30	VISNDVCAQV	10	0.74	1046.5
31	KLVVVGADGV	10	1.74	956.2
32	KLVVVGAAGV	10	1.83	912.0
33	NSFCMGGMNRR	11	2.83	1272.5
34	RPKPQQFFGLM	11	2.83	1348.4
35	ACLGRDRRTEE	11	3.72	1305.4
36	DAEKSDICTDEY	12	1.54	1387.5
37	TTIHYNYICNSS	12	1.83	1414.6
38	PHRERCSDSDGL-ace	12	3.64	1371.7
39	ACPGTDRRTGGGN	13	2.74	1261.4
40	ACPGKDRRTGGGN	13	3.74	1288.4
41	MGGMNWRPILTIIT	14	1.83	1603.0
42	SPALNKMFCELAKT	14	2.82	1552.7
43	HMTEVVRHCPHHER	14	6.32	1768.0
44	LAKTCPVRLWVDSTPP	16	2.74	1783.2
45	VVRRCPHQRCSDSDGI	16	4.65	1828.1
46	LGRNSFEVCVCACPGRD	17	2.73	1826.0
47	KLVVVGAGDVGKSALTI	17	2.74	1626.9
48	TPPPGTRVQQSQHMTEV	17	2.82	1893.0
49	YKLVVVGAAGVGKSALT	17	2.83	1632.0
50	YKLVVVGACGVGKSALT	17	2.83	1665.0
51	YNYMCNSSGMGGMNRRP	17	2.83	1938.5
52	YKLVVVGAVGVGKSALT	17	2.83	1661.0
53	YKLVVVGARGVGKSALT	17	3.83	1718.0
54	PPPGTRVRVMAIYKQSQ	17	3.83	1928.3
55	DGLAPPQHRIRVEGNLR	17	4.73	1928.2
56	VPYEPPEVGSVYHHPLQLHV	20	3.81	2297.6
57	FLTPKKLQCVDLHVISNDVCAQVHPQKVTK	30	6.65	3390.1
58	HQIINMWQEVGKAMYAPPISGQIRRIHIGPGRAFYTTKN	39	7.82	4481.2

voltage of 8000 V yielded $\mu_{eo} = 0.641 \cdot 10^{-5} cm^2\ V^{-1}\ s^{-1}$. From the above, an average value of $\mu_{eo} = 0.64 \cdot 10^{-5} \pm 0.005 \cdot 10^{-5}\ m^2\ V^{-1}\ s^{-1}$ was determined.

2.4. Determination of the electrophoretic mobility of a standard reference compound

Compound 1-4-dimethyl amino pyridine was selected as a reference, because it has a high electrophoretic mobility, it is chemically stable and is positively charged at pH 2.5. Using a 37-cm column and applying a voltage of 8 kV, an average migration time of 7.55 ± 0.13 was obtained from 25 separate injections over a period of two days. The apparent electrophoretic mobility of the standard $\mu_{app}(ref)$ was then determined according to the equation:

$$\mu_{app}(ref) = \frac{L.l}{Vt_m(ref)} = \frac{37 \cdot 30}{8000 \cdot 7.55 \cdot 60} = 30.6 \cdot 10^{-5}\ cm^{-2}\ V^{-1}\ s^{-1} \quad (5)$$

where L = total column length, l = injector-to-detector column length, V = applied voltage and $t_m(ref)$ = the migration time of the reference.

The electrophoretic mobility of the reference was determined according to the equation:

$$\mu_{ef}(ref) = \mu_{app}(ref) - \mu_{eo} = 30.62 \cdot 10^{-5} - 0.64 \cdot 10^{-5} = 29.98 \cdot 10^{-5} cm^2\ V^{-1}\ s^{-1} \quad (6)$$

In a separate determination using a 47-cm column, a value of $30.06 \cdot 10^{-5}\ cm^2\ V^{-1}\ s^{-1}$ was obtained for μ_{ef} (ref). The average of the two measurements ($30.02 \cdot 10^{-5}\ cm^2\ V^{-1}\ s^{-1}$) was used in the determination of peptide electrophoretic mobilities (as described in the next section) to correct for small day-to-day variations in μ_{eo}, if present.

2.5. Determination of the electrophoretic mobility of peptides

For each peptide, a 0.25–0.5-mg/ml solution was made, injected, and the migration time [t_m(pep)] recorded. Simultaneously, a sample of the reference was introduced as a second injection and its migration time was recorded, t_m(ref). The apparent electrophoretic mobility of each peptide [μ_{app}(pep)] was determined from the equation:

$$\mu_{app}(pep) = \mu_{app}(ref) \cdot \frac{t_m(ref)}{t_m(pep)} \quad (7)$$

Finally,

$$\mu_{ef}(pep) = \mu_{app}(pep) - \mu_{eo} \quad (8)$$

where $\mu_{eo} = \mu_{app}(ref) - 30.02 \cdot 10^{-5}$.

The μ_{ef} data are presented in Table 2.

2.6. Repeatability and reproducibility of μ_{ef} measurements

The electrophoretic mobility of six peptides and the reference standard was measured with each of two columns using two different buffer preparations on separate days. The results are presented in Table 3. Excellent repeatability was obtained with each column ($\pm 1\%$); however, the RSD of the averages from the two columns varied from 0.13 to 4.81, with an average of 2.34% for the set. This, we believe, represents the level of uncertainty of our experimental determination of electrophoretic mobilities.

3. Results and discussion

Despite the voluminous amount of work reported [3–14], there is still no consensus on the best approach for the correlation of peptide electrophoretic mobility with change and size parameters. Nyberg et al. [24] reported the first experiments to test the Offord model by CZE, and Rickard et al. [5] reported excellent correlation ($R^2 = 0.948$) of peptide μ_{ef} versus $q/M^{2/3}$ for 33 peptides and ten proteins. Grossman et al. [4], on the other hand, examined 40 peptides and obtained excellent correlation of μ_{ef} versus $[\ln (q+1)]/n^{0.43}$. Survay et al. [7] measured the μ_{ef} of oligoglycines and oligoalanines ($n = 2$–6) and tested the Offord model (μ_{ef} vs. $q/M^{2/3}$) and the Grossman et al. model [4,9,11] {μ_{ef} vs. $[\ln (q+1)/n^{0.43})]$} on their results and the results of Grossman et al. [4] and concluded that none is uniquely suitable for peptide mobility modeling. They based their conclusion on the fact that a wide range of exponents of the size parameters (M or n) can be

Table 2
Comparison of experimental electrophoretic mobility with theoretical calculation according to the equation $\mu_{ef} = (2.44 + 581.85\ q/M^{2/3}) \cdot 10^{-5} cm^2\ V^{-1}\ s^{-1}$

Peptide no.	$\mu_{ef} \times 10^5$ Experimental	$\mu_{ef} \times 10^5$ Theoretical	% Deviation[a]
1	10.31	12.02	−16.57
2	13.00	12.50	3.82
3	12.52	13.61	−8.74
4	21.70	21.13	2.61
5	19.27	18.88	2.02
6	18.43	18.12	1.71
7	15.39	15.90	−3.27
8	15.16	15.66	−3.33
9	14.86	15.13	−1.82
10	14.55	14.85	−2.07
11	13.90	14.22	−2.28
12	13.33	13.82	−3.68
13	13.86	13.76	0.71
14	12.81	12.98	−1.32
15	12.10	12.31	−1.76
16	11.05	11.52	−4.28
17	15.43	15.31	0.76
18	13.24	13.79	−4.17
19	10.38	10.59	−2.00
20	13.87	13.21	4.79
21	12.33	11.79	4.37
22	9.75	9.62	1.40
23	9.53	9.47	0.68
24	18.36	17.91	2.45
25	6.51	8.00	−22.90
26	6.02	6.67	−10.72
27	6.65	7.19	−8.10
28	19.71	18.41	6.60
29	16.56	17.14	−3.49
30	5.83	6.62	−13.58
31	13.13	12.88	1.94
32	14.1	13.77	2.36
33	18.3	16.47	10.02
34	16.98	15.94	6.14
35	20.97	20.57	1.93
36	9.91	9.65	2.65
37	10.59	10.89	−2.87
38	19.49	19.60	−0.57
39	15.08	16.10	−6.77
40	19.11	20.82	−8.97
41	10.2	10.22	−0.18
42	15.71	14.68	6.55
43	26.41	27.59	−4.48
44	15.12	13.29	12.13
45	21.46	20.55	4.25
46	13.66	13.08	4.27
47	13.69	13.97	−2.04
48	14.17	13.17	7.08
49	14.22	14.32	−0.73
50	14.34	14.17	1.21
51	14.3	13.04	8.84
52	15.06	14.20	5.73
53	17.8	17.98	−1.01
54	18.2	16.84	7.47
55	19.49	20.21	−3.69
56	15.32	15.18	0.94
57	18.68	19.59	−4.87
58	17.53	19.18	−9.44

[a] $\dfrac{\mu_{ef}(Exp) - \mu_{ef}(Theor)}{\mu_{ef}(Exp)} \cdot 100$

Table 3
Comparison of μ_{ef} data measured with two independent experimental set-ups[a]

Peptide	$\mu_{ef} \times 10^5$ cm^2 V^{-1} s^{-1}			
	Column 1	Column 2	Average±SD	RSD (%)
AA	19.00	19.53	19.27±0.27	1.37
AAA	15.80	15.00	15.43±0.40	2.60
YKLVVVGACGVGKSALT	13.82	14.86	14.34±0.52	3.62
VPYEPPEVGSVYHHPLQLHV	15.13	15.51	15.32±0.19	1.24
DGLAPPQHRIRVEGNLR	20.00	18.98	19.49±0.51	2.62
VLQELNVTV	6.33	6.97	6.65±0.32	4.81
1-4-Dimethyl amino pyridine (ref.)	30.06	29.98	30.02±0.04	0.13
			Average	2.34

[a] Column 1=47 cm and column 2=37 cm. Each data point in columns 1 and 2 is an average of three determinations to within±1% RSD.

used to fit the data with comparable precision. Conflicting results and conclusions were also reported by others [8–13]. Basak and Ladisch [12] and Cross and Cao [13] favored the Offord model; Hilser et al. [9] and Cifuentes and Poppe [11] favored the Grossman et al. model; while Gaus et al. [8] presented data that showed that neither the Offord model nor the Grossman et al. model is adequate. To add to the uncertainty, Chen et al. [10] presented peptide mobility data that gave excellent correlation with $q/M^{1/2}$.

The main objective of this study is to identify the empirical model that allows the most accurate prediction of electrophoretic mobilities of peptides from physical parameters. Once identified, the model is to be used for the practical goal of simulating peptide maps of proteins from protein amino acid sequences. For this purpose, a diverse set of peptides with wide ranges of charges and molar masses was assembled (Table 1), and their electrophoretic mobilities, at pH 2.5, were measured accurately (Table 2). As mentioned previously, pH 2.5 appears to be the most ideal for this application, because all peptides are positive at this pH and all move in the same direction towards the cathode. Also, the 10% polyacrylamide-coated columns are particularly suited for this application at pH 2.5, because (a) the coating is chemically stable and resists hydrolytic degradation, (b) it is hydrophilic, (c) it effectively shields the peptides from direct contact with the residual surface silanol groups, thus ensuring negligible electroosmotic flow, (d) it does not interact with the peptides by adsorption, or otherwise, thus ensuring high column efficiency and (e) the coating procedure yields reproducible columns with consistent day-to-day and batch-to-batch migration times (Table 3). Two figures are presented to illustrate the utility of this experimental set-up for the intended application. Fig. 1 shows the separation of a mixture of bioactive peptides and Fig. 2 shows the peptide map resulting from an endoproteinase Lys-C digest of β-lactoglobulin A. The separations presented above are reproducible over extended periods of time. We routinely use the same column for periods of months with no observable changes in performance, attesting to the ruggedness of this approach for this particular application. Only two columns were used to generate all of the data reported in this work.

3.1. Testing of existing models for peptide mobility correlations

The data reported in Table 2 was used to test the empirical model of peptide mobility in order to identify the model that most accurately reproduces our experimental data. Fig. 3 shows plots of peptide mobility versus (A) $q/M^{2/3}$, (B) $[Ln(q+1)]/n^{0.43}$, (C) $q/M^{1/3}$ and (D) $q/M^{1/2}$, respectively. Inspection of the plots and the R^2 values clearly demonstrated the superiority of the $q/M^{2/3}$ correlation. The equation of the best line from Fig. 3A was obtained by linear regression:

$$\mu_{ef} = (2.44 + 581.85\ q/M^{2/3}) \cdot 10^{-5} \text{cm}^2\ \text{V}^{-1}\ \text{s}^{-1} \quad (9)$$

Further evidence for the superiority of the Offord model comes from inspection and analysis of the

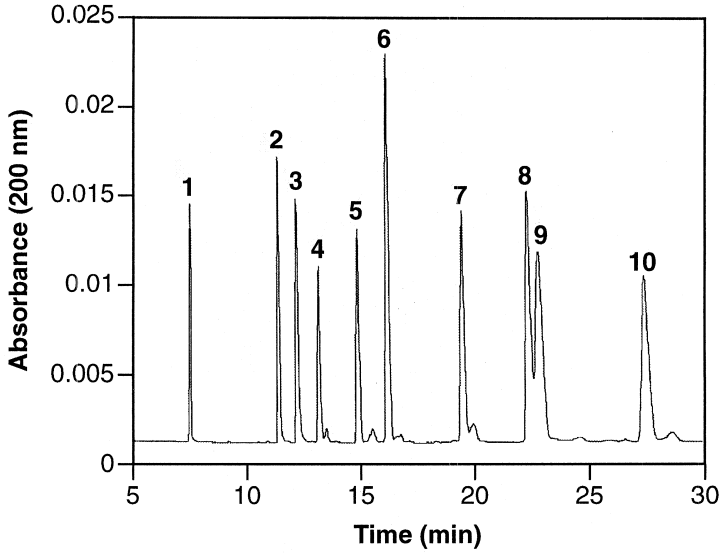

Fig. 1. Electropherogram showing the separation of a mixture of bioactive peptides at pH 2.5. Column: 10% T polyacrylamide-coated fused-silica; Column dimensions: 47 cm (effective length 40 cm)×50 μm I.D. Instrument: Beckman model P/ACE system 5500; voltage: 12 kV; current: 18 μA; injection: 5 s at 0.5 p.s.i.; buffer: 50 m*M* phosphoric acid adjusted to pH 2.5 with TEA; temperature: 22°C; detection: UV at 200 nm. Solutes: 1=reference; 2=bradykinin; 3=bradykinin fragment 1–5; 4=substance P; 5=[arg]–vasopressin; 6=luteinizing hormone releasing hormone; 7=bombesin; 8=leucine enkephalin; 9=methionine enkephalin; 10=oxytocin.

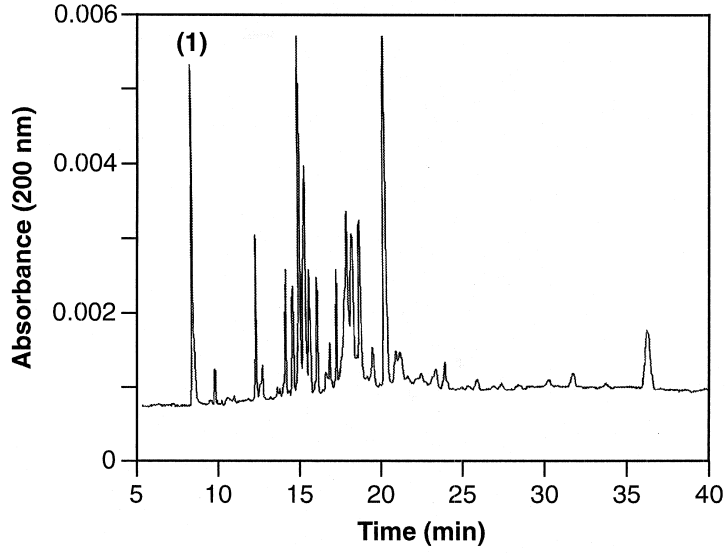

Fig. 2. Electropherogram showing the peptide map from endoproteinase Lys-C digestion of β-lactoglobulin A. Experimental conditions were the same as those given in Fig. 1. Digest: 100 μg of protein+10 μg of Lys-C in 100 μl of 25 m*M* Tris–HCl, pH 8.5 for 18 h at 37°C. Peak (1)=1-4-dimethylaminopyridine (Ref.).

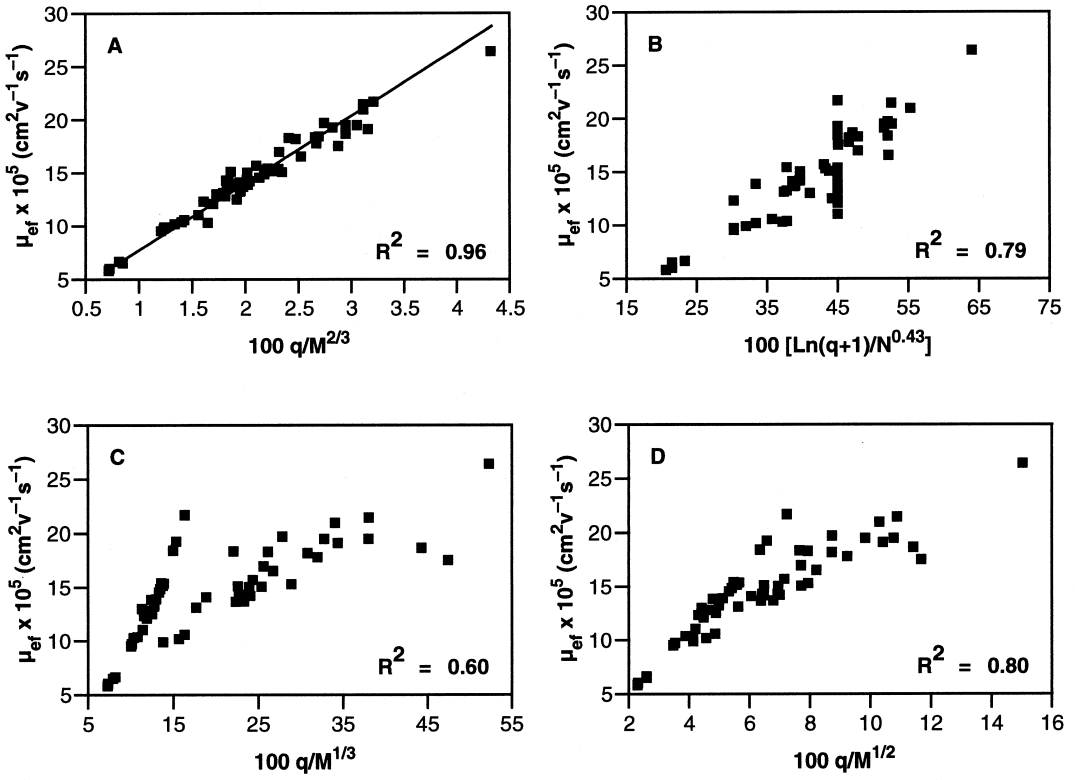

Fig. 3. Correlation of peptide electrophoretic mobility at pH 2.5 with charge-to-size parameters. Data are from Tables 1 and 2.

Table 4
Experimental electrophoretic mobilities for the monomers and dimers of cysteine-containing peptides

Peptide	Monomer $\mu_{ef} \times 10^5$ cm^2 V^{-1} s^{-1}	Dimer $\mu_{ef} \times 10^5$ cm^2 V^{-1} s^{-1}	$\frac{\mu_{ef}(\text{dimer})}{\mu_{ef}(\text{monomer})}$
LAKTCPVRLWVDSTPP	15.12	18.05	1.19
ACPGTDRRTGGGN	15.08	16.68	1.11
YNYMGNSSGMGGMNRRP	14.30	17.57	1.23
ACPGKDRRTGGGN	19.11	19.86	1.04
ACLGRDRRTEE	20.97	22.30	1.06
NSFCMGGMNRR	18.30	22.03	1.20
HMTEVVRHCPHHER	26.41	27.13	1.03
VISNDVCAQV	5.83	8.11	1.39
PHRERCSDSDGL-ace	19.49	22.38	1.15
TTIHYNYICNSS	10.59	13.80	1.30
DAEKSOICTDEY	9.91	12.63	1.27
YKLVVVGACGVGKSALT	14.34	17.92	1.25
		Average	1.19±0.11

data presented in Table 4. The table lists experimental μ_{ef} values for the monomers and sulfur-bridged dimers of 12 cysteine-containing peptides. With no exception, the dimers have higher μ_{ef} values. If one considers that dimerization results in doubling of charge and molar mass, then, according to the Offord model:

$$\frac{\mu_{ef}(\text{dimer})}{\mu_{ef}(\text{monomer})} = \frac{2q/(2M)^{2/3}}{q/M^{2/3}} = 1.26 \quad (10)$$

In contrast, $q/M^{1/3}$ and $q/M^{1/2}$ give ratios of 1.59 and 1.41, respectively. Inspection of the ratios presented in Table 4 reveals deviations from the theoretical value, which we mainly attribute to peptide shape and conformation effects that are not accounted for by the models.

In an effort to reconcile our results with others who obtained good results with the Grossman et al. model [4,9,11], and those who claimed no particular advantage of any model [8,10], we examined our experimental results in different ways.

Fig. 4 shows correlation of our experimental data according to the Grossman et al. model. Fig. 4A gives the correlation for all peptides and Fig. 4B gives the correlation for a subset of peptides with

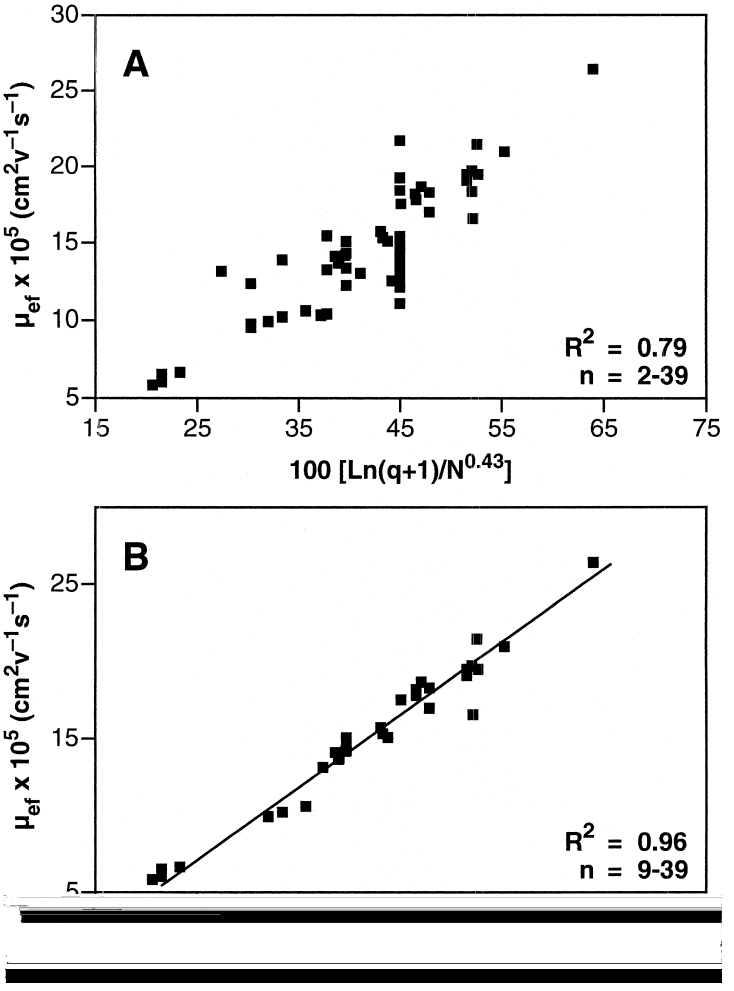

Fig. 4. Correlation of peptide electrophoretic mobility, at pH 2.5, according to the Grossman et al. model. (A) All peptides in Table 1; (B) selected range of peptides. Data are from Tables 1 and 2.

9–39 amino acid residues. Comparison of the two figures reveals that a dramatic improvement in the correlation is obtained when the small peptides ($n=$ 2–5) are excluded. This observation exposes an obvious deficiency in the Grossman et al. model, namely, it fails for peptides with the same number of amino acid residues and having comparable charges. Another interesting comparison is presented in Figs. 5 and 6, which show several correlation plots for subsets of peptides. Fig. 5 gives the correlations for peptides with five or less amino acid residues according to (A) $q/M^{2/3}$, (B) $q/M^{1/2}$ and (C) $q/M^{1/3}$, and Fig. 6 shows similar correlations for peptides with 9–17 amino acid residues. Examination of the R^2 values in both figures reveals no significant advantage of any of the models over the others. This result is in agreement with the observation of Survay et al. [7] and in apparent disagreement with the result of our correlation with the complete set of peptides (Fig. 3). This apparent contradiction could, however, be easily explained if one examines the mathematical properties of the plots of any of the functions $1/M^{2/3}$, $1/M^{1/2}$ or $1/M^{1/3.}$ versus one another. While curved over a wide range of M, the plots are, in fact, approximately linear over narrow ranges of M. Consequently, plots of peptide mobility versus any of the above functions will yield comparable R^2 values for subsets of peptides with narrow molar mass ranges. Differences in goodness-of-fit are only revealed when a large set of peptides with a wide molar mass range is considered, as demonstrated in Fig. 3.

Finally, Eq. 9 was used to generate the theoretical μ_{ef} values presented in Table 2. While most of the theoretical values fall within $\pm 5\%$ of the experimental values, there are numerous exceptions of larger divergence. However, examination of the data reveals no systematic deviation with respect to charge, molar mass or charge-to-molar mass ratio. A notable exception to this observation are small peptides with large positive charge, such as the lysine homologs with $n=2$–5, which systematically deviate from all models (data not shown). Thus, while the Offord model is superior to others, it still has room for improvement. As it stands, the model does not account for several factors that affect peptide mobility, such as peptide conformation and shape, peptide hydrophobicity, and shielding of charge in highly charged peptides [7,11,14,18]. In order to address the deficiencies inherent in the Offord model, we are currently investigating a new modeling approach that takes into account other physical properties that are neglected by current models. Each electrophoretic mobility data point is represented as a product of several functions representing charge, length, average mass, and position of charges relative to center of mass. This approach requires a large data-base of peptide mobilities of a diverse set of peptides that scan the above-mentioned parameters. Preliminary results are encouraging and a detailed report will be the subject of a future publication.

3.2. Theoretical simulation of peptide maps of protein digests

Peptide mapping by capillary electrophoresis [11,25–31] and capillary electrophoresis–mass spectrometry [25,32,33] is increasingly being utilized as a complement, if not a viable substitute, for the already established technique of high-performance liquid chromatography (HPLC). Capillary electrophoresis has several advantages over HPLC, including much higher resolution and a smaller sample size requirement. Moreover, capillary electrophoresis also offers a straightforward correlation of migration time with physical properties. The potential of using this property to simulate peptide maps of protein digests was explored in this work. Eq. 9 was used to simulate the theoretical peptide maps of a small peptide, melittin, and horse myoglobin and the resulting maps were compared with their experimental couterparts.

Fig. 7 gives a line representation of the endoproteinase Lys-C digest of the peptide sequencing standard: YAEGDVHATSKPARR. A 100-μg amount of the peptide was incubated for 18 h at 37°C with 10 μg of endoproteinase Lys-C in 100 μl of 25 m*M* Tris–HCl, pH 8.5. Under these conditions, endoproteinase Lys-C is known to specifically cleave peptide bonds at the C-terminal of lysine (K). The experimental electropherogram (data not shown) of the digest gave two peaks. Complete digestion was ensured by monitoring the disappearance of the parent peptide peak. The electrophoretic mobility of each peak was calculated from its migration time. The relative area percent was obtained from the

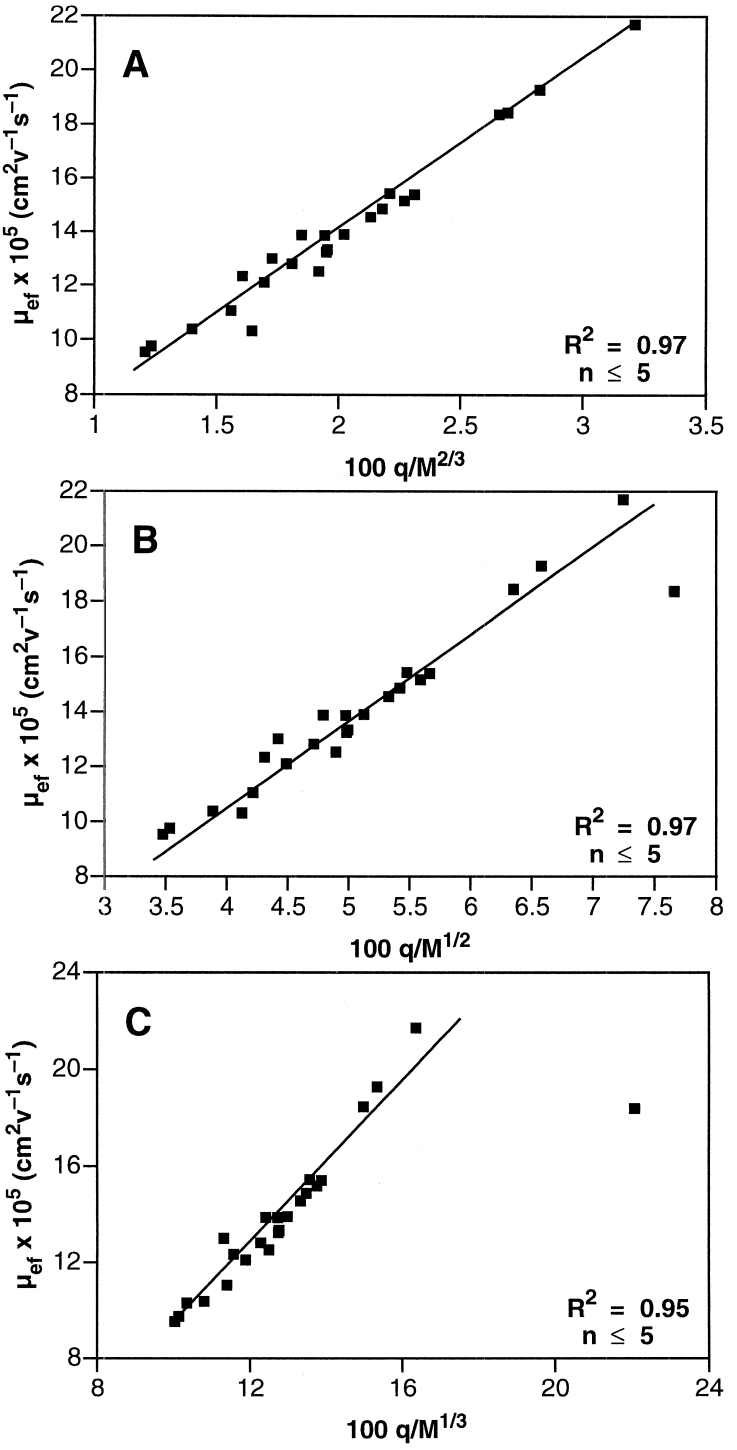

Fig. 5. Correlation of peptide electrophoretic mobility at pH 2.5 with charge-to-size parameters for peptides with five or less amino acid residues. Data are from Tables 1 and 2.

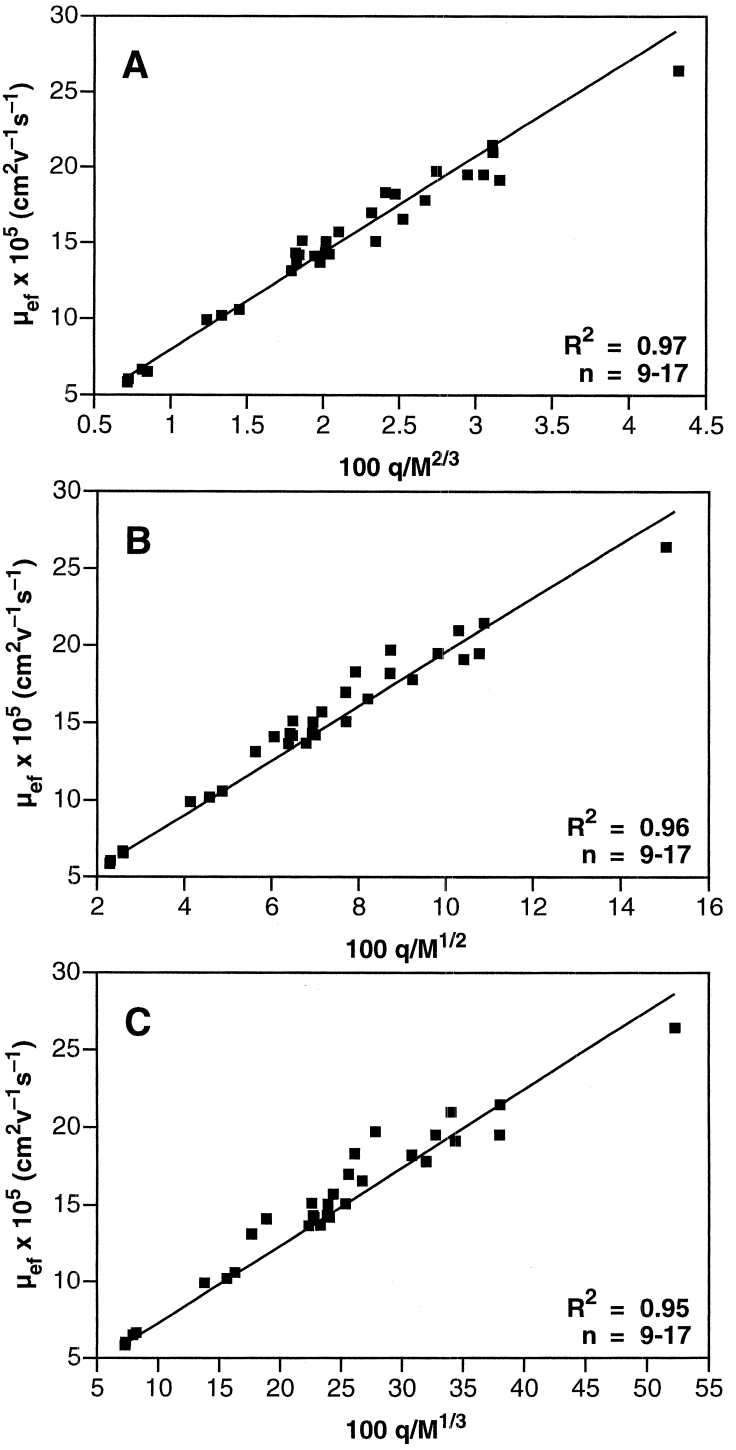

Fig. 6. Correlation of peptide electrophoretic mobility at pH 2.5 with charge-to-size parameters for peptides with 9–17 amino acid residues. Data are from Tables 1 and 2.

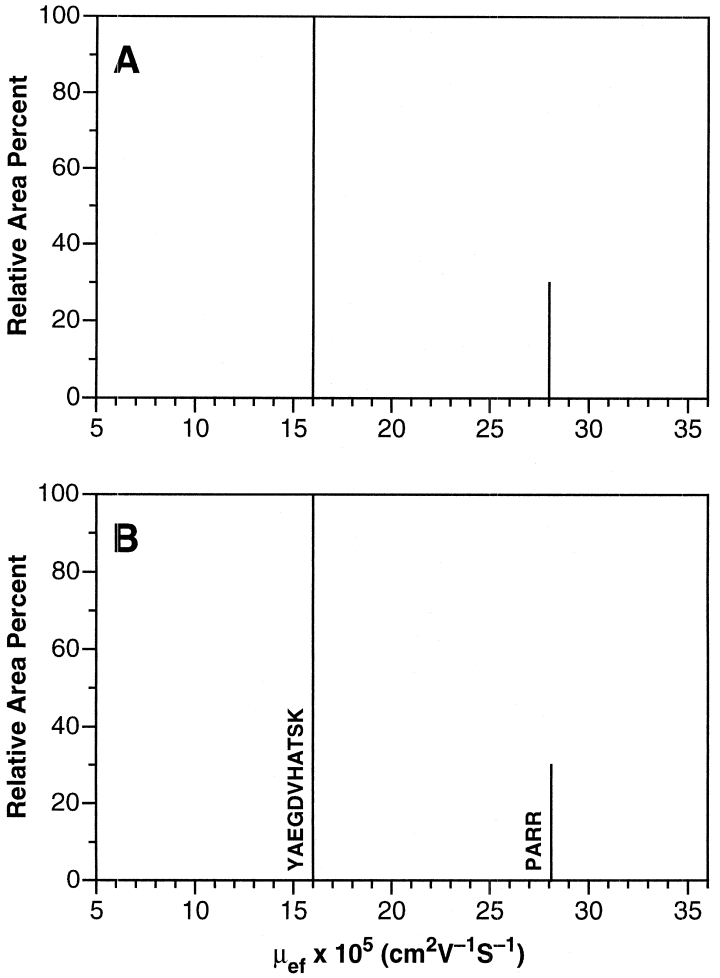

Fig. 7. Line representation of the mobility electropherogram of the endoproteinase Lys-C digest of the peptide sequencing standard: YAEGDVHATSKPARR. (A) Experimental conditions as in Fig. 1. Digestion as in Fig. 2. (B) Theoretical simulation.

corrected areas of the peaks. The larger peak was considered to be 100% and the smaller peak was normalized accordingly. The resulting line electropherogram is shown in Fig. 7A. The theoretical simulation was conducted as follows. The electrophoretic mobilities of the two fragments (YAEGDVHATSK and PARR) were calculated using Eq. 9. The relative area percent of each of the lines was considered to be proportional to the number of peptide bonds in the fragment. This is based on the assumption that the absorption of peptides at 200 nm (which is the wavelength used throughout this work to generate the experimental electropherograms) is proportional to the number of peptide bonds in the peptide. Experimental evidence supporting this assumption is presented by Herold et al. [34]. Here again, the absorbance of the larger fragment was considered to be 100%, and the absorbance of the smaller fragment was normalized accordingly.

The resulting line representation of the mobility electropherogram is given in Fig. 7B. Comparison of the two figures shows an excellent match between the experimental and the theoretical electropherograms, attesting to the viability of this approach.

Next, we tackled a slightly more complex mole-

cule. Melittin (GIGAVLKVLTTGLPALIS-WIKRKRQQ) was similarly treated with endoproteinase Lys-C. The resulting experimental electropherogram is shown in Fig. 8A and the corresponding theoretically simulated counterpart is shown in Fig. 8B. Here again, excellent matching of line position and relative area percent is obtained.

Finally, horse myoglobin ($n=153$) was similarly treated with endoproteinase Lys-C and the resulting electropherogram is given in Fig. 9A. The theoretical electropherogram, shown in Fig. 9B, was simulated as follows. First, the electrophoretic mobility of each fragment was calculated and then converted to migration time using the experimental parameters of column length and applied voltage, as was used to generate the experimental electropherogram. The peak shape was simulated by a Gaussian function, assuming that the peak area is proportional to $n-1$. The peak variance was arbitrarily adjusted so that it approximately matched the column efficiency (number of theoretical plates for the peak at 25 min) of its experimental counterpart. The agreement between the experimental and theoretical electropherograms, although not perfect, appears to be reasonably good, considering the deficiencies in the theoretical model and the possible problems associated with protein digestion reactions. Proteolysis reactions, in general, suffer from possible incomplete digestion, non-spe-

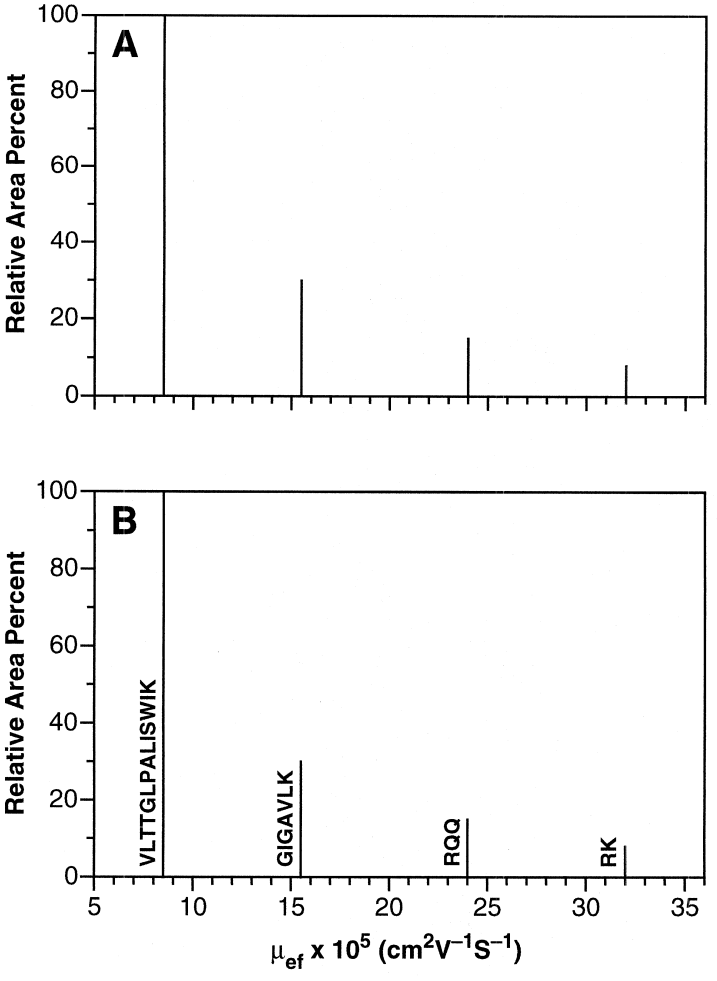

Fig. 8. Line representation of the mobility electropherogram of the endoproteinase Lys-C digest of melittin: GIGAVLKVLTTGLPALIS-WIKRKRQQ. (A) Experimental conditions as in Fig. 7. (B) Theoretical simulation.

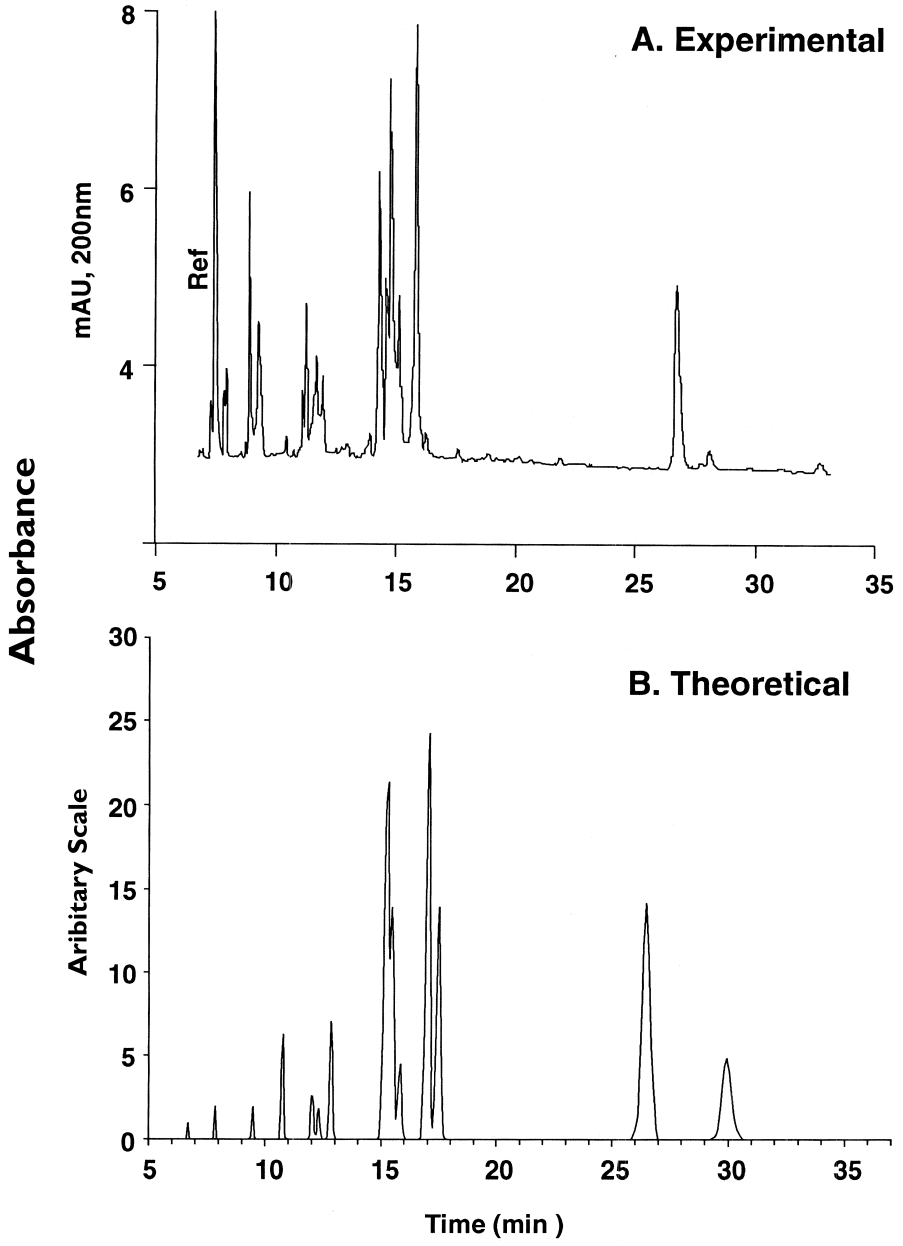

Fig. 9. Electropherogram of the endoproteinase digest of horse myoglobin. Experimental conditions as in Fig. 7.

cific digestion, and the possible presence of impurities in the protein starting material.

Finally, we believe that this approach is potentially useful for protein analysis and for the identification of the main fragments in peptide maps, especially if the protein digests of several proteolytic agents are considered simultaneously. Work along these lines is currently in progress in our laboratory.

Acknowledgements

The authors would like to thank Dr. J.A. Berzofsky, NCI, for supplying many of the peptides used in this study, and Dr. M. Kelsey, NCI, for helpful discussions.

This project has been funded in whole or in part with Federal funds from the National Cancer Institute, National Institutes of Health, under Contract No. NO-1-CO-56000.

The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products or organizations imply endorsement by the US Government.

References

[1] R.J. Wieme, in: E. Heftmann (Ed.), Chromatography, A Laboratory Handbook of Chromatographic and Electrophoretic Methods, 3rd ed, Van Nostrand Reinhold, New York, 1975, Ch. 10.
[2] P.D. Grossman, D.S. Soane, Anal. Chem. 62 (1990) 1592.
[3] P.D. Grossman, R.J. Wilson, G. Petrie, H.J. Lauer, Anal. Biochem. 173 (1988) 265.
[4] P.D. Grossman, J.C. Colburn, H.H. Lauer, Anal. Biochem. 179 (1989) 28.
[5] E.C. Rickard, M.M. Strohl, R.G. Nielsen, Anal. Biochem. 197 (1991) 197.
[6] B.J. Compton, J. Chromatogr. 559 (1991) 357.
[7] M.A. Survay, D.M. Goodall, S.A.C. Wren, R.C. Rowe, J. Chromatogr. 636 (1993) 81.
[8] H.-J. Gaus, A.G. Beck-Sickinger, E. Bayer, Anal. Chem. 65 (1993) 1399.
[9] V.J. Hilser Jr., G.D. Worosila, S.E. Rudnick, J. Chromatogr. 630 (1993) 329.
[10] N. Chen, L. Wang, Y.K. Zhang, Chromatographia 37 (1993) 429.
[11] A. Cifuentes, H. Poppe, J. Chromatogr. A 680 (1994) 321.
[12] S.K. Basak, M.R. Ladisch, Anal. Biochem. 226 (1995) 51.
[13] R.F. Cross, J. Cao, J. Chromatogr. A 786 (1997) 171.
[14] N.J. Adamson, E.C. Reynolds, J. Chromatogr. B 699 (1997) 133.
[15] R.E. Offord, Nature 211 (1966) 591.
[16] B. Skoog, A. Wichman, Trends Anal. Chem. 5 (1986) 82.
[17] C. Long (Ed.), Biochemist's Handbook, Van Nostrand, Princeton, NJ, 1961.
[18] M. Castagnola, L. Cassiano, I. Messana, G. Nocca, R. Rabino, D.V. Rossetti, B. Giardina, J. Chromatogr. B 656 (1994) 87.
[19] G.M. Janini, R.J. Fisher, L.E. Henderson, H.J. Issaq, J. Liq. Chromatogr. 18 (1995) 3617.
[20] R.M. McCormick, Anal. Chem. 60 (1998) 2322.
[21] J.B. Matthew, Annu. Rev. Biophys. Biophys. Chem. 14 (1985) 387.
[22] G.M. Janini, K.C. Chan, G.M. Muschik, H.J. Issaq, J. Chromatogr. A 657 (1994) 419.
[23] M.A. Hayes, I. Kheterpal, A.G. Ewing, Anal. Chem. 65 (1993) 27.
[24] F. Nyberg, M.D. Zhu, J.L. Liao, S. Hjerten, in: C. Schafer-Nielsen (Ed.), Proceedings of Electrophoresis 88, VCH, Weinheim, 1988, pp. 141–150.
[25] T. Van de Goor, A. Apffel, J. Chakel, W. Hancock, in: J.P. Landers (Ed.), Handbook of Capillary Electrophoresis, 2nd ed, CRC Press, Boca Raton, FL, 1997, pp. 213–258.
[26] M. Dong, R.P. Oda, M.A. Strausbauch, P.J. Wettstein, J.P. Landers, L.J. Miller, Electrophoresis 18 (1997) 1767.
[27] N. Bihoreau, C. Ramon, R. Vincentelli, J.-P. Levillain, D.V. Troalen, J. Cap. Electrophoresis 2 (1995) 197.
[28] M. Castagnola, L. Cassiano, R. Rabino, D.V. Rossetti, F.A. Bassi, J. Chromatogr. 572 (1991) 51.
[29] R.S. Rush, P.L. Derby, T.W. Strickland, M.F. Rohde, Anal. Chem. 65 (1993) 1834.
[30] T. Kornfelt, A. Vinther, G.N. Okafo, P. Camilleri, J. Chromatogr. A 726 (1996) 223.
[31] M. Stromquist, J. Chromatogr. A 667 (1994) 304.
[32] M.A. Winkler, S. Kundu, T.E. Robey, W.G. Robey, J. Chromatogr. A 744 (1996) 177.
[33] P. Cao, M. Moini, Rapid Commun. Mass Spectrom. 12 (1998) 864.
[34] M. Herold, G.A. Ross, R. Grimm, D.N. Heiger, in: K.D. Altria (Ed.), Capillary Electrophoresis Guidebook. Principles, Operation, and Application, Methods in Molecular Biology, Vol. 52, Humana Press, Totowa, NJ, 1996, pp. 285–308.