Journal of Chromatography, 572 (1991) 51–58
Biomedical Applications
Elsevier Science Publishers B.V., Amsterdam

CHROMBIO. 6087

Peptide mapping through the coupling of capillary electrophoresis and high-performance liquid chromatography: map prediction of the tryptic digest of myoglobin

MASSIMO CASTAGNOLA*, LOREDANA CASSIANO, RITA RABINO and DIANA VALERIA ROSSETTI

Istituto di Chimica, Facoltà di Medicina e Chirurgia "A. Gemelli", Università Cattolica (c/o Centro di Studio per la Chimica dei Recettori e delle Molecole Biologicamente Attive, C.N.R.), Largo F. Vito 1, 00168 Rome (Italy)

and

FRANCESCO ANDREASI BASSI

Istituto di Fisica, Facoltà di Medicina e Chirurgia "A. Gemelli", Università Cattolica, Largo F. Vito 1, 00168 Rome (Italy)

(First received May 8th, 1991; revised manuscript received July 30th, 1991)

ABSTRACT

The tryptic map of horse myoglobin was analysed through capillary electrophoresis using capillaries modified by a monolayer of acrylamide. The results were reproducible and the map was obtained in less than 30 min from *ca.* 8 pmol of tryptic digest. The peptide identification was performed using peptides previously identified by high-performance liquid chromatography. The peak areas measured using the two techniques are closely related, and the comparison of clution and migration times shows that the two techniques provide different maps. Furthermore, using the semiempirical relationship suggested by Grossman *et al.* [Anal. Biochem., 179 (1989) 28], which links the electrophoretic mobility to the charge of the peptide and its number of amino acids, a good agreement between predicted and experimental mobilities was observed.

INTRODUCTION

Peptide mapping after selective endoprotease digestion plays a central role in any protein sequence analysis; it is the first step in any strategy for a sequence. In addition, by comparative analysis it allows information to be obtained about protein variants and post-translational protein modifications. Furthermore, a good correspondence between experimental and predicted selective endoprotease mapping might be used as a control of genetic engineering peptide products.

Because of its high sensitivity and high resolving power, we use capillary electrophoresis (CE) for the analysis of tryptic mapping. Moreover, using a semiempirical relationship between peptide charge and dimension and its mobility, similar to that proposed by Grossman *et al.* [1], a satisfactory agreement between

0378-4347/91/$03.50

theoretical and experimental mobilities seems to be possible. Therefore the theoretical mapping obtained from a tryptic digest of myoglobin was compared with the experimental one obtained by CE; as a means of checking the mapping identification, the results obtained by reversed-phase high-performance liquid chromatography (HPLC) were employed.

EXPERIMENTAL

Apparatus

The CE apparatus was a P/ACE 2000 (Beckman, Palo Alto, CA, USA) connected to P/ACE software. HPLC was performed on a Millipore-Waters (Milford, MA, USA) chromatographic station equipped with two 510 pumps, a Wisp 712 automatic injector, a Lambda-Max 481 LC spectrophotometer and Baseline 810 software.

Materials

All common reagents were analytical grade, purchased from Merck (Darmstadt, Germany) or Farmitalia-Carlo Erba (Milan, Italy). Horse myoglobin (skeletal muscle) was purchased from Calbiochem (San Diego, CA, USA), trypsin (bovine pancreas) TPCK-treated, was from Sigma (St. Louis, MO, USA), phenyl isothiocyanate and triethylamine were from Pierce (Rockford, IL, USA), acrylamide (ultrapure) was from USB (Cleveland, OH, USA) and acetonitrile (HPLC grade) was from Carlo Erba (Milan, Italy). The buffers for CE and HPLC were filtered on Flowpore FN (0.2 μm) (Flow Labs., McLean, VG, USA).

Procedures

Trypsinization of myoglobin. Horse myoglobin was converted into the apoprotein form after haem removal following the acetone–acid method of Rossi-Fanelli *et al.* [2]. The trypsinization was performed as previously described [3], and the tryptic digest was lyophilysed.

Capillary electrophoresis. CE was performed on capillaries modified following the procedure of Hjerten [4]. The capillary was 57.5 cm long (50.0 cm at the detection window), with an inner diameter of 75 μm assembled in a Beckman cartridge. The buffer used was 80 m*M* sodium phosphate (pH 2.50). The applied voltage was fixed at 25 kV with an initial current of *ca.* 110 μA. The temperature of the analysis was always fixed at 20°C. The lyophilysed tryptic digest was dissolved in the buffer at a concentration of 15 mg/ml. The sample was injected by pressure, and the measured injection volume was *ca.* 8 nl, corresponding to *ca.* 8 pmol of myoglobin tryptic digest. The wavelength of detection was 214 nm. After each analysis, a 3-min washing step with filtered, bidistilled water was performed.

Reversed-phase HPLC. A Brownlee (Santa Clara, CA, USA) Aquapore RP300 octyl (7 μm) 220 mm × 4.6 mm I.D. cartridge, protected with 15 mm × 4.6 mm I.D. guard cartridge of the same resin, was used for reversed-phase

HPLC. Eluent A was the same buffer used for CE, and eluent B was acetonitrile–eluent A (60:40, v/v). The gradient applied was linear from 0 to 60% of B in 60 min. The myoglobin tryptic digest was at a concentration of 60 mg/ml, and the injection volume was 2 μl in analytical runs and 20 μl in semipreparative runs. The wavelength was 214 nm. In semipreparative runs the peaks were collected, dried under vacuum, and dissolved in bidistilled water.

Amino acid analysis. The peptides were hydrolysed by azeotropic HCl in the gas phase using a PicoTag Workstation (Waters), and analysed as phenylthiocarbamyl derivatives with a PicoTag column used following the specification of manufacturer.

RESULTS AND DISCUSSION

Initial peptide separation by CE on unmodified capillaries gave non-reproducible results: the measured migration time variability was greater than 8%, and resolution parameters [electrophoretic theoretical plate number (N) and resolution] were very poor. The presence of a high electroosmotic flow (f_{eo}) and a strong interaction of the peptides with the inner wall of the capillary were the presumed explanation. Because the use of furfuryl alcohol as an f_{eo} marker instead of acetanilide, also at low pH (*i.e.* pH 2.5) when the ionization of silanolic groups is very low, provided a lower migration time, the interaction should be ascribed not only to charge interaction but also to hydrogen bonding between the peptide and the silanolic groups of the surface.

Following the suggestion of Bushey and Jorgenson [5], the use of lysine in the separation buffer, which competes with the peptide for the wall interaction, provided better resolution parameters, but the observed reproducibility was always unsatisfactory. For these reasons the capillary was modified by a monolayer of acrylamide, following the procedure suggested by Hjerten [4].

Using this precaution a very reproducible peptide map was observed (variation of migration times less than 2%) with very high resolution parameters (mean measured N *ca.* 200 000). Owing to the decreased interaction between the peptides and the capillary wall, a very large decrease of f_{eo} was observed. In fact, the use of neutral markers for the f_{eo} measure was impossible because no response within an acceptable time was observed. In addition, the tryptic maps in sodium or potassium phosphate buffers were identical, although in the presence of f_{eo} the substitution of sodium with potassium reduces f_{eo} itself (data not reported).

Fig. 1 shows the tryptic mapping obtained from 8 pmol of horse myoglobin digest. Owing to the difficulties connected with the collection and amino acid analysis of peptides, the peptide identification was performed through the coupling of the HPLC mapping shown in Fig. 2, achieved using the same buffer and the same wavelength (214 nm) used for CE separation. Each peptide identified and separated by HPLC was added to the whole myoglobin digest and analysed by CE. After peak identification, a correlation between the peak areas obtained

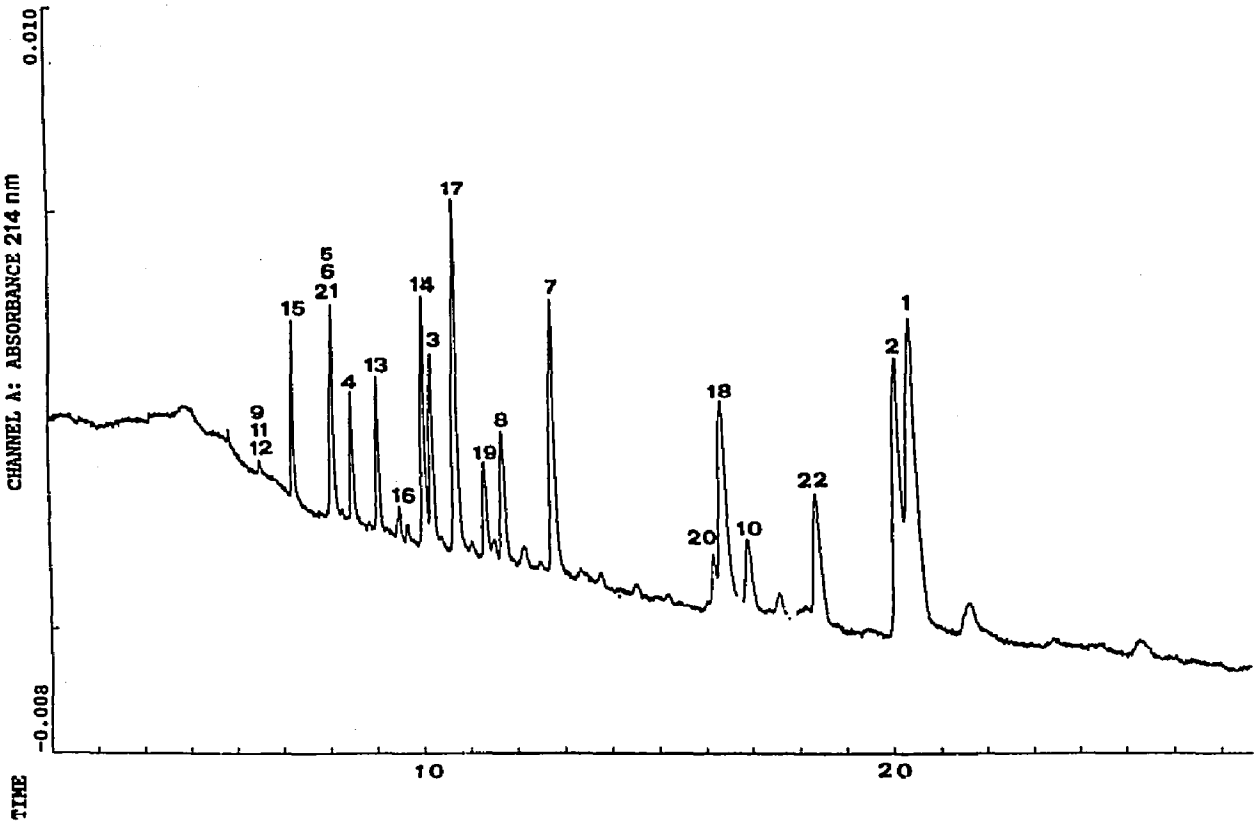

Fig. 1. Horse myoglobin tryptic mapping obtained from 8 pmol of whole digest by CE on a modified capillary. The electrophoretic conditions are described in Experimental.

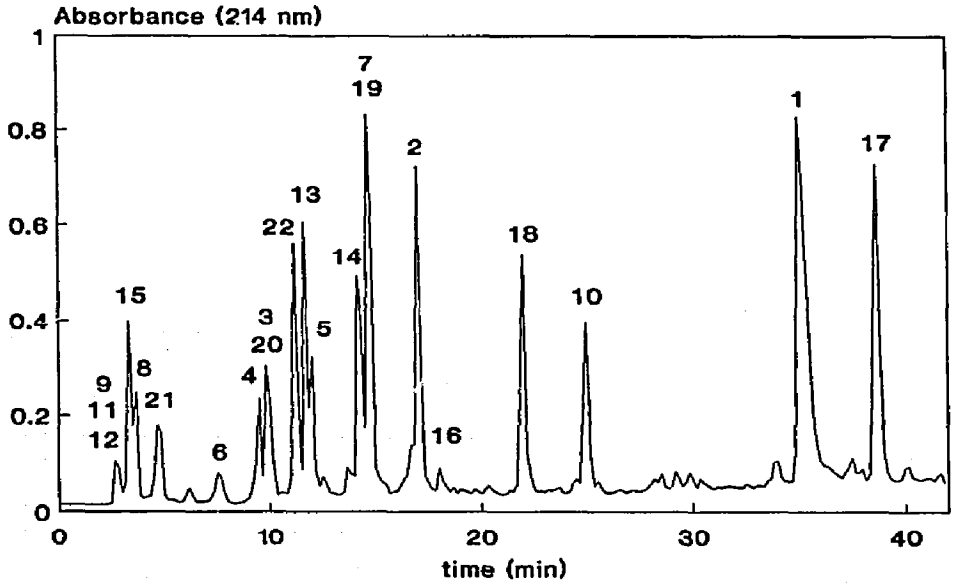

Fig. 2. Horse myoglobin tryptic mapping obtained by reversed-phase HPLC. The chromatographic conditions are described in Experimental.

by HPLC and CE was possible. To obtain this correlation a correction on the peak areas obtained by CE was necessary. Whereas in HPLC any substance crosses the spectrophotometric cell at a constant speed carried by a constant elution flow, the crossing time of a substance separated by CE is inversely proportional to its mobility; thus the peak width increases as a function of the time of migration, beyond the normal increase due to the dispersion factors. For these reasons, after the integration of the peaks obtained by CE, they were normalized by a modification of the formula proposed by Huang *et al.* [6]:

$$A_{corr} = A_m(l/t) - 1.26hd_r \quad (1)$$

where A_m is the measured area, t is the migration time of the peak, l is the length of the capillary from the injection space to the spectrophotometric window, h is the peak height and d_r is the linear vertical dimension of the window (in our system 200 μm). The area obtained after this correction is expressed in spatial dimensions instead of temporal dimensions; therefore, for a comparison with HPLC results, the percentage of each CE peak area obtained after the correction was related to the percentage of the appropriate HPLC peak area. The results of the linear correlation after this area comparison are indicated in Fig. 3. The good correlation ($r = 0.976$) ensures that a direct identification is possible by comparing HPLC and CE peak areas.

Moreover, the comparison between the HPLC elution times and the CE migration times for each tryptic peptide, as shown in Fig. 4 and Table I, demonstrates that the two techniques cannot be considered orthogonal from the statisti-

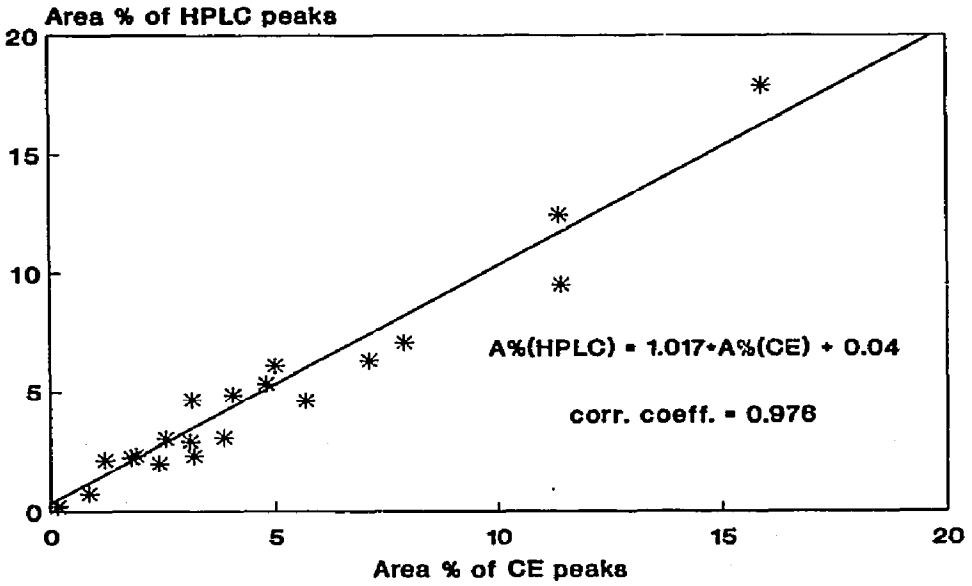

Fig. 3. Linear correlation obtained from a comparison of the measured percentage area of HPLC peaks and the percentage areas measured on CE mapping and corrected through eqn. 1.

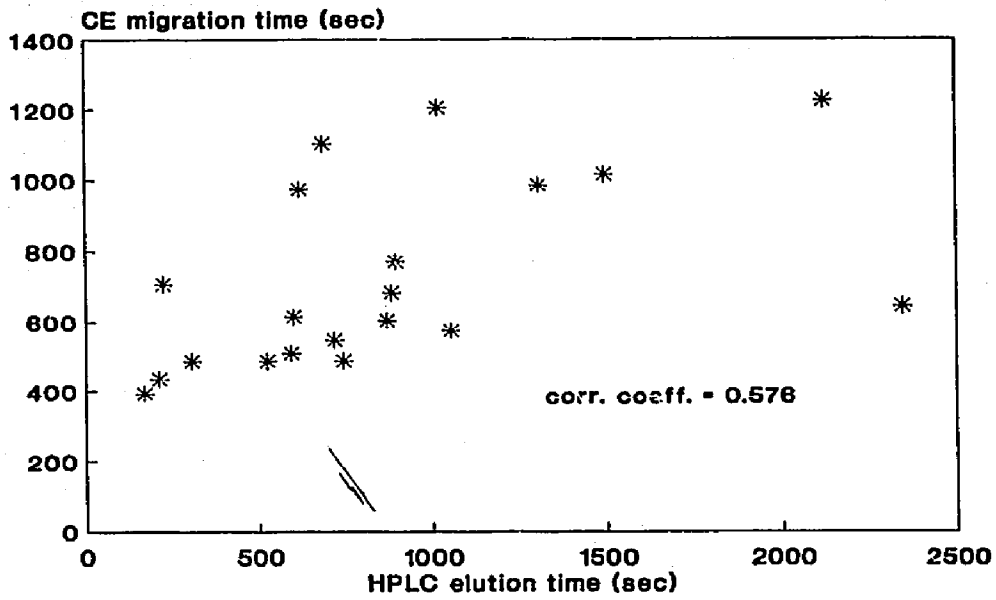

Fig. 4. Correlation between HPLC elution times and CE migration times of the peptides identified in the tryptic digest of horse myoglobin.

cal point of view ($r = 0.576$). Nevertheless, both techniques are useful as they give complementary results. The HPLC separation derives from different peptide polarities, whereas electrophoretic separation is connected to the peptide charge.

These experimental results point to the value of a possible theoretical prediction of the peptide mobility, for the purpose of using CE endoprotease mapping as a control of the proteins produced by genetic engineering techniques, and as a tool for fast, sensible and accurate analysis of protein variants and modifications.

The theory of electrophoresis indicates that the mobility of a substance is given by:

$$\mu = q/(6\pi\eta r) \tag{2}$$

where η is the viscosity of the solution, q is the charge and r is the molecular radius of the substance. Grossman *et al.* [1] have shown that this equation is not appropriate for peptides, and have suggested the semiempirical relationship.

$$\mu = A\ln(q+1)/n^B + C \tag{3}$$

where A, B and C are three parameters obtained by a best fit from their experimental results, and n is the number of amino acids in the peptide; in particular, the value of A is governed by the system used.

To verify the agreement of their equation with the tryptic pattern of horse myoglobin, the peptide charge at pH 2.5 was computed following a modification of the program suggested by Skoog and Wichman [7]: the result is shown in Table I. Then a multi-parameter best fit to the equation, using the measured mobility

TABLE I

CHROMATOGRAPHIC AND ELECTROPHORETIC DATA OF TRYPTIC PEPTIDES OBTAINED FROM HORSE MYOGLOBIN

Tryptic peptide	Peptide sequence[a]	Elution time HPLC (s)	Migration time CE (s)	Mobility observed ($\times 10^4$) ($cm^2/V\ s$)	Charge[b]	Mobility predicted[c] ($\times 10^4$) ($cm^2/V\ s$)
1	GLSDGEWQQVLNVWGK	2121	1225	0.93	1.28	1.1
2	VDADIAGHGQEVLIR	1020	1205	0.96	2.24	1.4
3	LFTGHPETLEK	602	613	1.87	2.30	1.6
4	FDK	594	507	2.26	1.30	1.9
5	FK	744	485	2.37	1.33	2.3
6	HLK	524	485	2.37	2.33	2.5
7	TEAEMK	897	768	1.49	1.30	1.5
8	ASEDLK	225	704	1.63	1.28	1.5
9	K	169	392	2.93	1.33	2.9
10	HGTVVLTALGGILK	1495	1016	1.13	2.33	1.4
11	K	169	392	2.93	1.33	2.9
12	K	169	392	2.93	1.33	2.9
13	GHHEAELK	718	544	2.11	3.30	2.0
14	PLAQSHATK	870	602	1.91	2.33	1.7
15	HK	211	434	2.64	2.33	2.9
16	IPIK	1053	573	2.00	1.33	1.8
17	YLEFISDAIIHVLHSK	2345	643	1.78	3.28	1.6
18	HPGNFGADAQGAMTK	1306	985	1.16	2.30	1.4
19	ALELFR	884	680	1.69	1.32	1.5
20	NDIAAK	620	972	1.28	1.30	1.5
21	YK	306	485	2.37	1.33	2.2
22	ELGFQG	688	1102	1.00	0.43	1.0

[a] One-letter code.

[b] The charge was computed at pH 2.50 following a modification of the program of Skoog and Wichman [7].

[c] The predicted mobility was computed according to eqn. 2 and the parameters listed in Table II.

and the Marquardt [8] algorithm, was performed. The parameters obtained are shown in Table II, in which the values obtained by Grossman *et al.* [1] are also reported for comparison.

The higher A value obtained by Grossman *et al.* [1] probably arises from the use of buffers of different ionic strength and capillaries of different dimensions. The value of B obtained by our fit indicates that probably the mobility of a peptide is inversely proportional to the square root of the number of amino acids in the chain. This observation is in good agreement to a random coil peptide conformation; in fact, it is well known that the molecular radius of a polymer in a random coil conformation is proportional to the square root of the number of freely rotating bonds. In effect, in the previous form of their equation, Grossman *et al.* [1] proposed the use of the square root of the number of amino acids; after the fit with their peptide mobilities they changed the value to 0.43. Again, the

TABLE II

PARAMETERS OBTAINED FOR EQN. 2

Parameter	Marquardt algorithm [8]	Grossman *et al.* [1]
A	$(2.78 \pm 0.27) \cdot 10^{-4}$	$5.23 \cdot 10^{-4}$
B	0.51 ± 0.08	0.43
C	$(0.59 \pm 0.23) \cdot 10^{-4}$	$0.25 \cdot 10^{-4}$
	$r = 0.936$	

difference between the two fits could derive from the use of different capillaries and buffers.

Therefore the following relationship, with only two parameters, between peptide mobility and its charge can be proposed:

$$\mu = A\ln(q+1)/\sqrt{n} + C \quad (4)$$

The satisfactory correlation obtained ($r = 0.936$) between theoretical and experimental mobilities demonstrates that the equation is a good tool for obtaining an acceptable prediction of CE peptide mobilities.

CONCLUSION

CE is a very powerful technique for the analysis of an endoprotease mapping and it provides separation which is complementary to that obtainable by HPLC methods. Bushey and Jorgenson [9] have demonstrated that the coupling of the two techniques is technically achievable and provides well resolved two-dimensional maps. The advantages of CE derive from its very low sample requirement and the good agreement between calculated and measured mobilities; using reversed-phase HPLC mapping the same high correlation is not obtainable [10]. The disadvantages of CE include the difficulty of fraction collection, which indicates the need for appropriate standards or combination with HPLC separation for correct peptide identification.

REFERENCES

1 P. D. Grossman, J. C. Colburn and H. H. Lauer, *Anal. Biochem.*, 179 (1989) 28.
2 A. Rossi Fanelli, E. Antonini and A. Caputo, *Biochim. Biophys. Acta*, 30 (1958) 608.
3 M. Castagnola, L. Cassiano, R. De Cristofaro, S. Luciani, D. V. Rossetti and R. Landolfi, *J. Chromatogr.*, 494 (1989) 310.
4 S. Hjerten, *J. Chromatogr.*, 347 (1985) 191.
5 M. M. Bushey and J. W. Jorgenson, *J. Chromatogr.*, 480 (1989) 301.
6 X. Huang, W. F. Coleman and R. N. Zare, *J. Chromatogr.*, 480 (1989) 95.
7 B. Skoog and A. Wichman, *Trends Anal. Chem.*, 5 (1986) 82.
8 D. W. Marquardt, *J. Soc. Ind. Appl. Math.*, 11 (1963) 431.
9 M. M. Bushey and J. W. Jorgenson, *Anal. Chem.*, 62 (1990) 978.
10 M. Castagnola, L. Cassiano, R. De Cristofaro, R. Landolfi, D. V. Rossetti and G. B. Marini Bettolo, *J. Chromatogr.*, 440 (1988) 231.