ELSEVIER

Journal of Chromatography A, 667 (1994) 304–310

JOURNAL OF CHROMATOGRAPHY A

Short Communication

Peptide mapping using combinations of size-exclusion chromatography, reversed-phase chromatography and capillary electrophoresis

Mats Strömqvist

Department of Protein Chemistry, Symbicom AB, P.O. Box 1451, S-901 24 Umeå, Sweden

(First received August 19th, 1993; revised manuscript received February 14th, 1994)

Abstract

Recombinant extracellular superoxide dismutase was proteolytically degraded by trypsin and the digest was thereafter separated using three different separation techniques. The size differences of the obtained fragments were exploited by size-exclusion chromatography (SEC) on a highly efficient column for peptide separation. Collected fractions representing different size groups of peptides were then separated on the basis of hydrophobicity on reversed-phase liquid chromatography (RPLC). This led to simplified RPLC-chromatograms where collected peaks were considerably more pure than if the digest was separated direct on RPLC and the identity of eluted peaks could easily be determined by amino acid analysis. Finally the digest was run on capillary electrophoresis for separation based on charge differences. For the identification of the different peaks in the electropherogram, the total digest was spiked with known peptides purified by the two other separation techniques. This technique proved to be very powerful for peak identification in an electropherogram from a digest composed of many fragments eluting closely together.

1. Introduction

Among the methods available for confirming the primary structure of proteins, peptide mapping is probably the most powerful. For recombinant proteins to be used as drugs it has become the most essential analysis for confirming identity between different batches to ascertain the quality of the product [1].

The protein is chopped into smaller pieces by means of a chemical or an enzymatic cleavage. The obtained peptides are thereafter separated on the basis of different properties such as size, charge and hydrophobicity (for a review see ref. 2). Since the introduction of peptide mapping the separation technique has been developed from slab-gel electrophoresis and thin-layer chromatography [3,4] to high-performance liquid chromatography [5–7] and capillary electrophoresis (CE) [8].

The outstanding method in recent years has been reversed-phase liquid chromatography (RPLC) separation which is an excellent method to detect small changes within a protein: a single amino acid difference in a peptide often gives rise to a change in the elution volume of that particular fragment. From the digest of a small protein with no or a few modifications, such as glycosylation, most peptides are usually well separated by RPLC and can be identified by amino acid analysis or mass spectrometry. However, when a protein is larger and more complex,

0021-9673/94/$07.00
SSDI 0021-9673(94)00145-Y

a single dimension analysis is usually not sufficient for a good exploration of the primary structure [7]. If the digest contains a large number of hydrophilic fragments, these might not be bound to the column and therefore elute unretarded and thus not separated. The option in that case is to use several different cleavage methods and/or use several different separation techniques to increase the knowledge of the structure.

Through the breakthrough of CE during the past five to six years (reviewed in refs. 9 and 10), this technique has been introduced and established as an excellent way of separating protein digests [11–13]. CE has the advantage of being very fast and the separating power is close to that of RPLC [14]. CE has fruitfully been combined with RPLC [15–17] and with SEC [18] to increase the information gained from peptide mapping.

In this paper it is shown how three different separation techniques separating on the basis of the size, hydrophobicity and charge of the generated peptides can be combined to reveal as much as possible of the structure of recombinant human extracellular superoxide dismutase (EC-SOD).

2. Experimental

2.1. Cleavage of EC-SOD

Recombinant EC-SOD produced in Chinese hamster ovary (CHO) cells was purified to more than 98% purity, carboxymethylated and cleaved with trypsin as described earlier [19].

2.2. Reversed-phase liquid chromatography

An Ultrasphere microbore C_{18} (250 × 2.0 mm I.D., Beckman Instruments, Palo Alto, CA, USA) coupled to a Beckman System Gold chromatography unit was used for all samples. The eluent system was 0.1% trifluoroacetic acid (TFA) in water–acetonitrile, the flow rate was 0.2 ml/min and temperature control (38°C) was achieved using a Waters (Milford, MA, USA) TCM temperature control unit.

2.3. Size-exclusion chromatography

For size separation of the peptides a polyhydroxyethyl aspartamide (200 × 9.4 mm I.D., 200 Å, Poly-lc Inc., Columbia, MD, USA) column was used. The buffer system was 0.2 *M* sodium sulfate–5 m*M* potassium phosphate, pH 3.0, containing 25% acetonitrile. The flow rate was 1 ml/min, the absorbance at 214 nm was recorded and the chromatography was performed at 22°C.

2.4. Capillary electrophoresis

A PACE 2100 (Beckman) was used for capillary electrophoresis. Samples were diluted in 10 m*M* H_3PO_4, pH 2.5, injected to the capillary (fused silica, 100 μm × 50 cm) by 10-s pressure injection, separated in 0.1 *M* H_3PO_4, pH 2.5 at 20 kV. Absorbance data at 214 nm was collected.

2.5. Amino acid analysis

The amino acid content was analyzed by PITC-derivatization and separation on an Ultrasphere ODS column (150 × 4.6 mm I.D., Beckman) as described earlier [19].

3. Results and discussion

Deduced from the amino acid sequence [20], EC-SOD should theoretically give 26 fragments upon cleavage with trypsin (Table 1). However, many of the fragments are too short to be identified by measuring the peptide bond at 214 nm. The composition of the digest is also complicated by the six adjacent cleavage sites at the C-terminus of the protein (T19–T24) and that some unspecific cleavage also occurs. Taken together this implies that it is difficult to separate all fragments by only one chromatographic step.

Reversed-phase liquid chromatography of the digest produces a chromatogram as shown in Fig. 1a. Many of the generated peptides could be identified in the chromatogram but most col-

Table 1

Theoretical tryptic fragments of EC-SOD

Fragment	Amino acid sequence
T1	WTGEDSAEPNSDSAEWIR
T2	DMYAK
T3	VTEIWQEVMQR
T4	R
T5	DDDGTLHAACQVQPSATLDAAQPR
T6	VTGVVLFR
T7	QLAPR
T8	AK
T9	LDAFFALEGFPTEPNSSSR
T10	AIHVHQFGDLSQGCESTGPHYNPLAVPHPQHPGDFGNFAVR
T11	DGSLWR
T12	YR
T13	AGLAASLAGPHSIVGR
T14	AVVVHAGEDDLGR
T15	GGNQASVENGNAGR
T16	R
T17	LACCVVGVCGPGLWER
T18	QAR
T19	EHSER
T20	K
T21	K
T22	R
T23	R
T24	R
T25	ESGCK
T26	AA

The fragment numbers start at the N-terminus of the intact protein.

lected fractions contained several fragments and the identity was therefore difficult to determine by total amino acid analysis. Amino acid composition analysis of the unretarded fraction from RPLC showed a high content of arginine which is not surprising since short fragments containing arginine could be expected both from the cluster of cleavage sites at the C-terminal part and from short internal fragments.

To improve the fingerprint, the digest was subjected to size-exclusion chromatography. The development of columns separating also short peptides with high resolution has improved this technique considerably as a tool for peptide mapping. Fig. 1b shows the result of SEC of the digest. The first peak eluted was a single fragment much larger than the other ones and was easily identified as the only glycosylated fragment, T9 plus an incompletely cleaved variant of it, T8 + 9. No other fragments could be directly identified from the size separation.

The separation of the EC-SOD digest on capillary electrophoresis (CE) is shown in Fig. 1c. The resolution is higher than in the SEC run, most fragments are well separated. No attempt to collect different fractions were made since the fragments eluted closely together in the electropherogram and were therefore difficult to collect as pure peaks.

To simplify the identification of the peaks within the RPLC-chromatogram and the CE-electropherogram, the different separation techniques were combined. First, the digest was separated on the basis of the fragment size on the SEC column. Six different fractions were collected from the SEC separation (as indicated

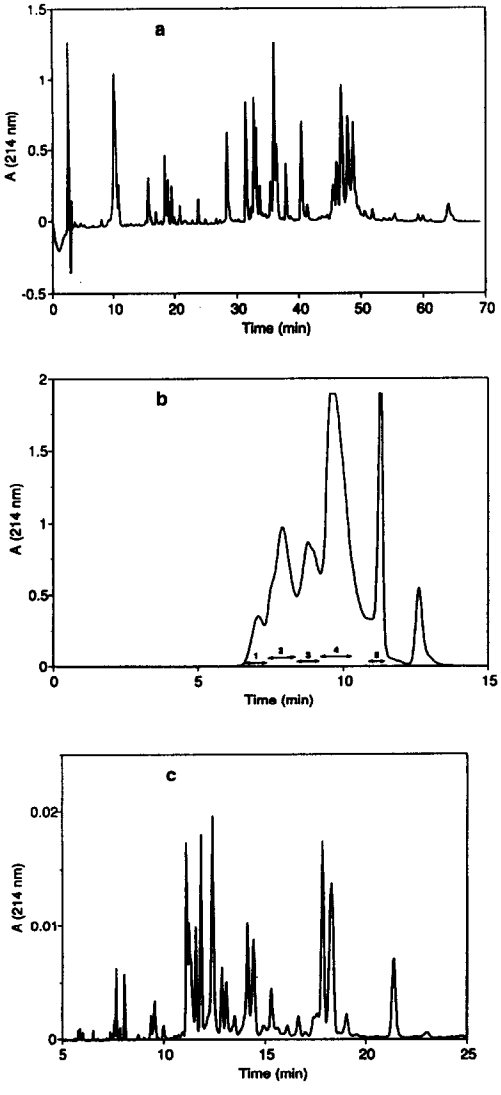

Fig. 1. (a) Tryptic digest of EC-SOD separated using reversed-phase chromatography. The gradient program started at 0.1% trifluoroacetic acid (TFA) in water, a linear gradient of 0–24% acetonitrile in 0.1% TFA from 5–30 min followed by a linear gradient from 24–36% acetonitrile over the next 30 min. The column was thereafter washed by a steep gradient up to 60% acetonitrile in 0.1% TFA and a 10-min wash at that acetonitrille concentration before re-equilibration to the starting conditions. The flow rate was 0.25 ml/min and 230 μg EC-SOD digest was loaded onto the column. (b) Size-exclusion chromatography of 400 μg EC-SOD digest, and (c) capillary electrophoresis of 10 μg EC-SOD digest. Fractions 1–5 collected from the SEC in (b) were further analyzed by RPLC (Fig. 2). The conditions for the chromatography in (b) and the electrophoresis in (c) are described under Experimental.

in Fig. 1b) and thereafter separated on the RPLC column (Fig. 2). The collected fractions from those RPLC runs were easily identified in most cases by amino acid composition analysis. Much help in the identification was gained from the information where this fragment eluted on SEC since this gives a useful hint on the size of the fragment.

Since there is no baseline resolution of the collected peaks from the SEC run, some fragments also appeared in the neighbouring peaks collected. After this two-dimensional separation all collected peaks were more than 80% pure according to analysis by CE (not shown) and their identity could be determined by amino acid analysis. When fractions were collected from a RPLC separation of the digest before it was separated on SEC, the dominating fragment in some fractions constituted only approximately 40% of the material even though the peak shape indicated that it was pure (not shown).

To identify the peaks in the electropherogram in Fig. 1c, the total digest was spiked with the purified and identified peaks from SEC–RPLC. Since the peaks of the electropherogram were so close together, running the purified individual fragments separately on CE gave not a certain identification. The amount of spiking fraction was chosen to give about twice the peak height of the particular peak. The spiking fraction was dissolved in 0.1% TFA in water and did not exceed 10% of the total sample volume when mixed with the total digest. The identity of the peak could therefore be obtained in only one run. Some of the spiked samples are shown in Fig. 3.

The best method for identification of peaks in an electropherogram is to use a mass spectrometer coupled to the CE instrument [21]. Alternatively peaks can be collected when eluted and thereafter subjected to mass spectrometry or amino acid sequencing. Collection techniques are dependent on a good resolution of the peaks because of a delay period between the time when the peak is detected and the time when it elutes from the capillary end. Another drawback with the collection technique is that the amount of digest that can be applied to the capillary at each

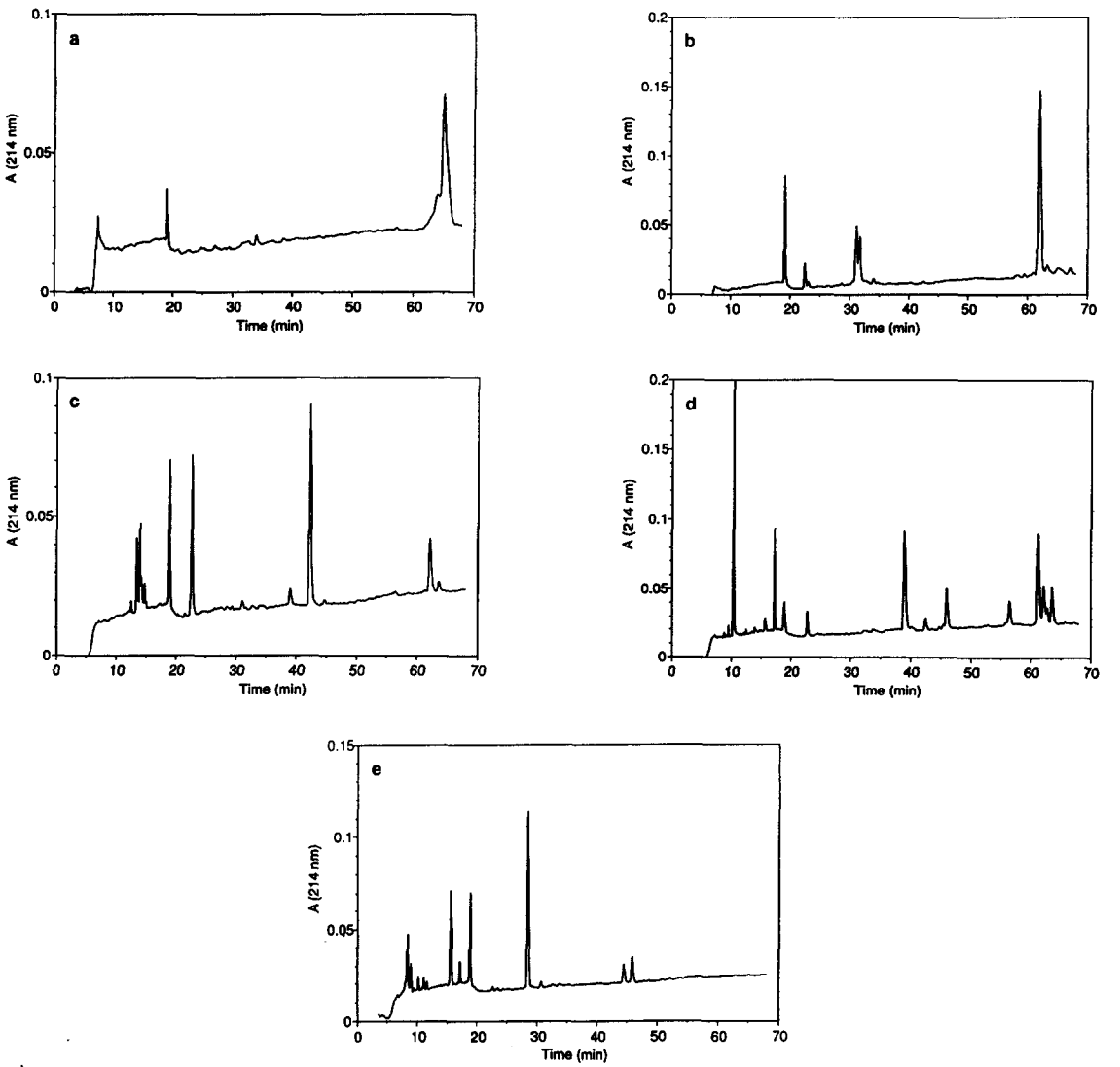

Fig. 2. RPLC of collected fractions 1–5 (a–e) from SEC of the tryptic digest of EC-SOD (Fig. 1b). Fractions collected in Fig. 1b were concentrated in a speed-vac and 20–50% of the pool was loaded onto the column. The gradient started by a linear increase from 0–15% acetonitrile in water in the first 15 min followed by a linear increase from 15–27% acetonitrile from 15–65 min and thereafter a steep increase up to 60% acetonitrile in one min. The column was thereafter washed and re-equilibrated. Both the water and the acetonitrile contained 0.1% TFA.

run is restricted. Therefore, it may be necessary to run the sample several times to collect amounts sufficient for analysis. The spiking technique used here is in many respects superior to the collection technique. The amount needed for this identification was minimal. This is a very powerful method that has also been used by others for the identification of peaks in electropherograms [22,23]. Fractions eluting very closely together can be identified, something which is very difficult otherwise, even if the reproducibility of the system is high. Since the

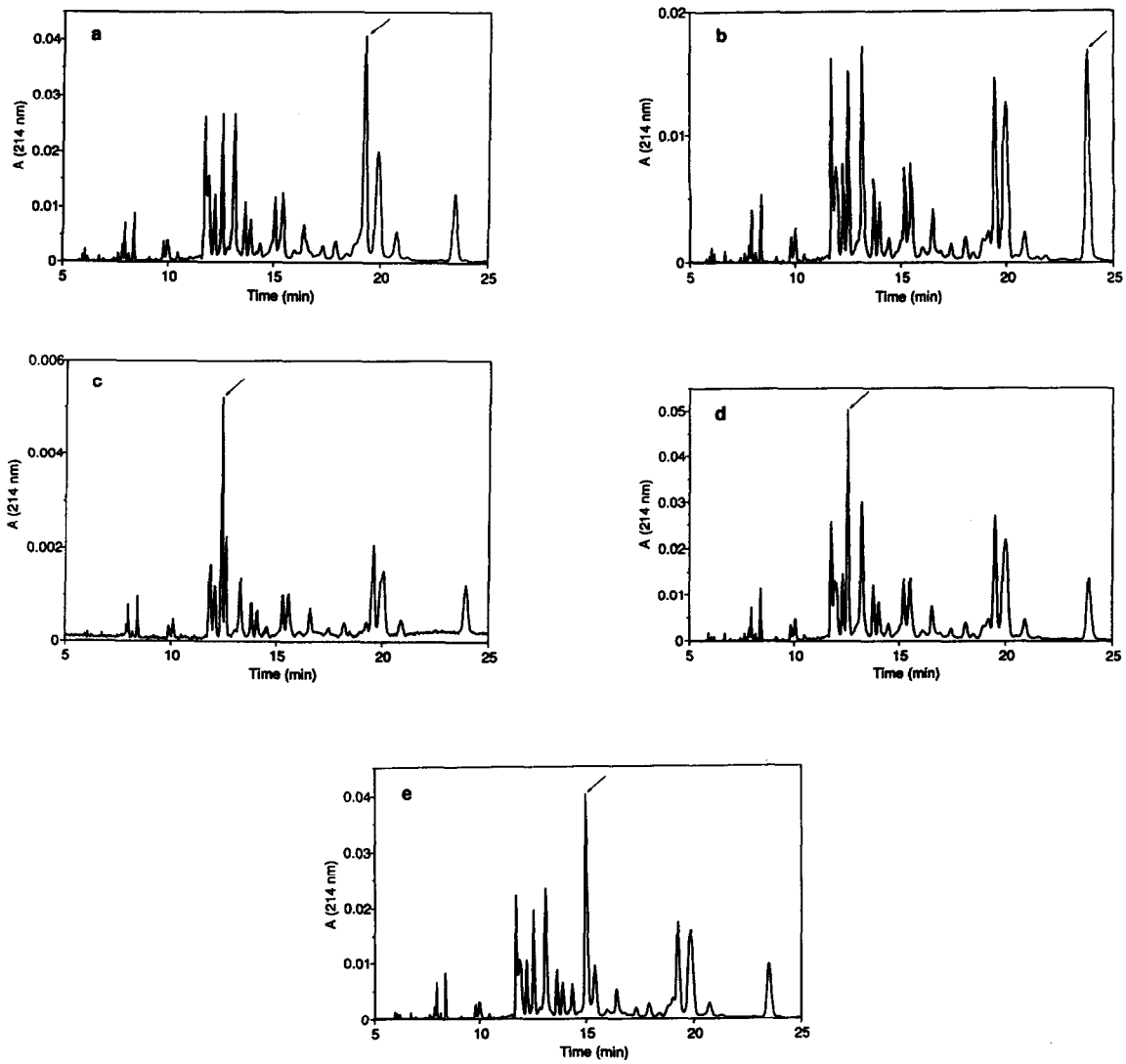

Fig. 3. Capillary electrophoresis of the tryptic digest of EC-SOD spiked with some of the identified peptides collected from the reversed-phase chromatographies described in Fig. 2. The EC-SOD digest (2–10 μg) was spiked with an amount of the purified peptide to obtain approximately twice the peak height of that particular peak in the chromatogram. The elution of the spiked peptide is indicated with an arrow. The digest was spiked with (a) fragment T1, (b) fragment T3, (c) fragment T6, (d) fragment T11 and (e) fragment T15.

particular fragment that is used for spiking is mixed with the total digest before electrophoresis, comigration is a very convincing evidence for identity, especially when only major fragments of the digest are used for spiking. In Fig. 3 the variation in the elution also can be seen. The elution time of the individual peaks varies too much between runs to make it possible to collect peaks with accuracy despite that the system was well equilibrated. By using the spiking technique

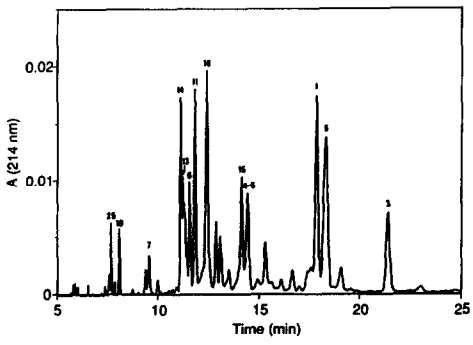

Fig. 4. The capillary electropherogram earlier shown in Fig. 1c with all identified peptides marked. The peptides were identified using the technique shown in Fig. 3 and are named according to the T-numbers they are given in Table 1.

this problem was overcome and most of the peptides could be identified in the electropherogram (Fig. 4).

4. Acknowledgements

The author acknowledges the excellent technical assistance of Helen Fält and Thord Johansson as well as the financial support of Symbicom AB.

5. References

[1] R.L. Garnick, N.J. Solli and P.A. Papa, *Anal. Chem.*, 60 (1988) 2546.
[2] E.A. Carrey, in T.E. Creighton (Editor), *Protein structure: a practical approach*, IRL Press, Oxford, 1989, p. 117.
[3] D.W. Cleveland, S.G. Fischer, M.W. Kirschner and U.K. Laemmli, *J. Biol. Chem.*, 252 (1977) 1102.
[4] R.E. Stephens, *Anal. Biochem.*, 84 (1978) 116.
[5] W.S. Hancock, C.A. Bishop, R.L. Prestidge and M.T.W. Hearn, *Anal. Biochem.*, 89 (1978) 203.
[6] F.E. Regnier, *LC·GC*, 5 (1987) 100.
[7] F.E. Regnier, *LC·GC*, 5 (1987) 230.
[8] J.W. Jorgenson and K.D. Lukacs, *J. High Resolut. Chromatogr. Chromatogr. Commun.*, 4 (1981) 230.
[9] Z. Deyl and R. Struzinsky, *J. Chromatogr.*, 569 (1991) 63.
[10] W.G. Kuhr and C.A. Monnig, *Anal. Chem.*, 64 (1992) 389R.
[11] K.A. Cobb and M. Novotny, *Anal. Chem.*, 61 (1989) 2226.
[12] R.G. Nielsen and E.C. Rickard, *J. Chromatogr.*, 516 (1990) 99.
[13] K.A. Cobb and M.V. Novotny, *Anal. Chem.*, 64 (1992) 876.
[14] J. Frenz, S.L. Wu and W.S. Hancock, *J.Chromatogr.*, 480 (1989) 379.
[15] M.M. Bushey and J.W. Jorgenson, *Anal. Chem.*, 62 (1990) 161.
[16] M.M. Bushey and J.W. Jorgenson, *Anal. Chem.*, 62 (1990) 978.
[17] J.P. Larmann Jr., A.V. Lemmo, A. W. Moore Jr. and J.W. Jorgenson, *Electrophoresis*, 14 (1993) 439.
[18] H. Yamamoto, T. Manabe and T. Okuyama, *J. Chromatogr.*, 515 (1990) 659.
[19] M. Strömqvist, J. Holgersson and B. Samuelsson, *J. Chromatogr.*, 548 (1991) 293.
[20] K. Hjalmarsson, S.L. Marklund, Å. Engström and T. Edlund, *Proc. Natl. Acad. Sci. U.S.A.*, 84 (1987) 6340.
[21] W.M.A. Niessen, U.R. Tjaden and J. van der Greef, *J. Chromatogr.*, 636 (1993) 3.
[22] T.E. Wheat, P.M. Young and N.E. Astephen, *J. Liq. Chromatogr.*,14 (1991) 987.
[23] M. Castagnola, L. Cassiano, R. Rabino, D.V. Rossetti and F.A. Bassi, *J. Chromatogr.*, 572 (1991) 51.