Journal of Chromatography, 627 (1992) 137–143
Elsevier Science Publishers B.V., Amsterdam

CHROM. 24 586

Use of monobromobimane to resolve two recombinant proteins by reversed-phase high-performance liquid chromatography based on their cysteine content

Donald O. O'Keefe

Department of Analytical Research, Merck Research Laboratories, Rahway, NJ 07065 (USA)

Ann L. Lee and Shigeko Yamazaki

Department of Biochemical Process R & D, Merck Research Laboratories, West Point, PA 19486 (USA)

(Received July 2nd, 1992)

ABSTRACT

A rapid reversed-phase HPLC assay useful for fermentation and downstream process development was developed for monitoring transforming growth factor-α-*Pseudomonas aeruginosa* exotoxin A 40 (TGFα-PE40). This protein is a chimeric recombinant protein synthesized in *Escherichia coli*. In the fermentation, full-length TGFα-PE40 is present along with PE40, an $M_r \approx 40\,000$ C-terminal fragment of TGFα-PE40, which co-purifies with TGFα-PE40 in many cases. A highly efficient reversed-phase HPLC assay using ultraviolet absorbance detection provided excellent resolution of the chimeric protein from the host-cell proteins in the crude cell lysate. However, this technique failed to resolve TGFα-PE40 from PE40, thereby limiting its use for in-process quantitation of the product. In order to resolve these two proteins, we have developed a new technique based on the sulfhydryl specificity of the fluorescent probe monobromobimane. Treatment of in-process samples with dithiothreitol followed by monobromobimane produces fluorescently-labeled TGFα-PE40, but does not label PE40 due to the lack of cysteine residues in this fragment. Thus, reversed-phase HPLC analysis using fluorescence detection provides the selectivity necessary to discriminate between TGFα-PE40 and PE40.

INTRODUCTION

Transforming growth factor-α-*Pseudomonas aeruginosa* exotoxin A 40 (TGFα-PE40, $M_r = 44\,960$) is a recombinant fusion protein produced in *Escherichia coli* [1,2]. TGFα-PE40 is a potential anticancer agent that is cytotoxic towards human tumor cells expressing specific growth factor receptors [2]. The N-terminal portion of the protein is composed of transforming growth factor-α (TGFα, 50 amino acids) and is responsible for binding to specific cell surface receptors. The C-terminal portion of the molecule is a fragment of *Pseudomonas aeruginosa* exotoxin A (PE) called PE40 ($M_r \approx 40\,000$) [3]. PE40 contains a domain that translocates the protein's catalytic subunit across cell membranes allowing inactivation of cellular protein synthesis.

To facilitate process development of TGFα-PE40, a rapid assay was needed that could monitor the product at each step of the fermentation and purification process. Process monitoring and quantitation of TGFα-PE40, especially during fermentation, was a challenging endeavor for two main reasons. TGFα-PE40 was an intracellular product which in the early stages of development represented only a small percentage of the total cellular pro-

Correspondence to: D. O. O'Keefe, Department of Analytical Research, Merck Research Laboratories, Merck and Co., Inc., P.O. Box 2000, 80M-112, Rahway, NJ 07065, USA.

0021-9673/92/$05.00

tein. This made monitoring fermentation samples difficult. Nevertheless, despite the crude nature and the low product level of fermentation and early purification samples, a highly efficient reversed-phase HPLC assay was developed that provided excellent resolution from the host-cell proteins. However, it appeared that the synthesis of TGFα-PE40 was accompanied by the production of a related molecule. Electrophoretic and western blot analysis suggested that this molecule was PE40. Although TGFα-PE40 and PE40 differ by approximately 50 amino acids, the reversed-phase HPLC assay with UV absorbance monitoring failed to resolve the two proteins. This led to inaccurate quantitation of fermentation expression levels of the product. Additionally, the assay lost its usefulness in estimating the recovery and yield of TGFα-PE40 at each step in the purification process. Although TGFα-PE40 and PE40 were resolved by western blot analysis, this laborious and low-throughput technique was impractical for process monitoring.

To differentiate PE40 from TGFα-PE40, the 6 cysteine residues exclusively located in the TGFα domain were exploited by labeling them with the sulfhydryl-specific reagent monobromobimane (4-bromomethyl-3,6,7-trimethyl-1,5-diazabicyclo-[3.3.0]octa-3,6-diene-2,8-dione; mBBr) [4]. As a result, TGFα-PE40 was fluorescently labeled and PE40 was not. The fluorescent-labeling did not affect the retention time of TGFα-PE40 so the two products still co-eluted upon reversed-phase analysis. However, by using fluorescence detection and UV absorbance monitoring, the necessary selectivity to discriminate between TGFα-PE40 and PE40 was obtained. This technique provided a more quantitative and accurate method for monitoring TGFα-PE40 during fermentation and downstream processing steps.

EXPERIMENTAL

Materials

Trifluoroacetic acid (TFA) and dithiothreitol (DTT) were obtained from Pierce (Rockford, IL, USA). Tris base was purchased from Boehringer Mannheim (Indianapolis, IN, USA). Acrylamide was acquired from National Diagnostics (Manville, NJ, USA). Other electrophoresis reagents and rabbit anti-goat antibody conjugated to alkaline phosphatase were purchased from Bio-Rad (Richmond, CA, USA). HPLC-grade acetonitrile was acquired from Fisher. Other reagents were obtained from either Fisher or Sigma.

Western blots

Protein samples were electrophoresed through 12% polyacrylamide–sodium dodecyl sulfate (SDS) gels according to the method of Laemmli [5]. The samples did not contain reducing agent and were not heated prior to electrophoresis. After electrophoresis, proteins were transferred to nitrocellulose paper (Schleicher & Schuell, Keene, NH, USA) in a Genie electroblotter (Idea Scientific, Minneapolis, MN, USA) as described by Towbin *et al.* [6] except the transfer buffer was made 0.1% SDS. Proteins bound to the nitrocellulose paper were then probed with either goat anti-TGFα antisera (Biotope, Redmond, WA, USA) or goat anti-PE antisera (List Biologicals, Campbell, CA, USA) as described [7]. Immunoreactive proteins were then detected using rabbit anti-goat antibody conjugated to alkaline phosphatase [8,9]. Control experiments demonstrated that these antisera did not cross react with *E. coli* proteins.

Monobromobimane labeling

To expose cysteine sulfhydryls, purified protein and in-process samples were denatured and reduced in 200 m*M* Tris–HCl, pH 8.0, 1% SDS, 3 m*M* EDTA, 3 m*M* DTT for 20 min at room temperature. mBBr (Molecular Probes, Eugene, OR, USA) was added to a final concentration of 15 m*M* from a 100 m*M* stock solution in acetonitrile. The reaction proceeded for 2 min at room temperature in subdued light before it was terminated by a 10-fold dilution into 4 m*M* cysteine. Cysteine and DTT stock solutions were made fresh.

Chromatography

Proteins were chromatographed on a HY-TACH non-porous C_{18} column (30 × 4.6 mm I.D.) from Glycotech with a linear gradient of 34 to 64% acetonitrile in 0.1% TFA over 6 min at 1.0 ml/min. The column was equipped with a water jacket equilibrated at 80°C by a Lauda RM6 circulating water bath. Control experiments demonstrated that TGFα-PE40 remained intact during the short time it was exposed to elevated temperature. UV absor-

bance was at 280 nm and fluorescence detection was at 470 nm after excitation at 382 nm. The chromatography system consisted of a Waters Model 712 WISP autosampler and a Waters Model 680 automated gradient controller controlling Waters Model 510 and Model M-6000 pumps. A Spectroflow 757 variable-wavelength absorbance detector from Applied Biosystems and a Spectro Vision FD-200 fluorescence detector were in-line to monitor column effluent.

RESULTS AND DISCUSSION

An in-process HPLC assay was developed to monitor TGFα-PE40 levels at each step of the fermentation and purification process. This rapid, reversed-phase HPLC assay employed a short column (30 × 4.6 mm I.D.) packed with a micropellicular, silica-based octadecyl stationary phase operated at elevated temperature and with high mobile phase flow-rate. Under these conditions, which facilitated rapid and highly efficient separations and facile column regeneration [10], the column provided the resolving power required to separate TGFα-PE40 from host-cell proteins in crude cell lysates. For the analysis of crude cell lysates from fermentation broths, samples were mixed with 6 *M* guanidine–HCl, acidified, centrifuged and then injected in the HPLC system. As shown in the upper panel of Fig. 1, TGFα-PE40 was only a minor component in that particular cell lysate. Comparison with the bottom panel indicated that the identified peak is likely to represent the product since its retention time corresponded with the purified reference material. In addition, there was little interference in the 5-min region from *E. coli* proteins determined by analyzing a cell lysate lacking TGFα-PE40. It is also noted that the total time of analysis was only 8 min and column regeneration was completed within 2 min.

During process development, SDS–polyacrylamide gel electrophoresis of TGFα-PE40 samples showed the continual presence of an $M_r \approx 40\,000$ protein. Samples enriched in TGFα-PE40 or the $M_r \approx 40\,000$ protein were prepared and chromatographed on the HY-TACH non-porous C_{18} column using a slightly modified linear gradient of acetonitrile from the conditions described in Fig. 1 (see Experimental). It was found that the two proteins could not be resolved and that they had a retention time of 4.24 min (Fig. 2). Analysis of the samples by

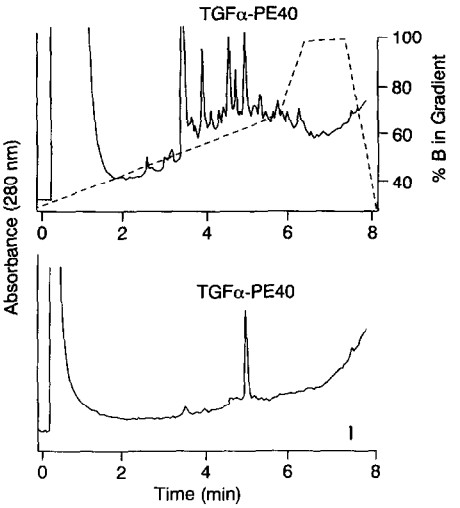

Fig. 1. Rapid reversed-phase HPLC chromatograms of a crude cell lysate and TGFα-PE40. Conditions: Glycotech HY-TACH C_{18} micropellicular (non-porous) silica-based column (30 × 4.6 mm I.D.). Gradient: A: 0.1% TFA in water; B: 0.1% TFA in acetonitrile–water (80:20); temperature, 80°C; flow-rate, 2 ml/min; 0.05 AUFS. Top panel: a crude cell lysate containing TGFα-PE40 was mixed with 6 *M* guanidine HCl, acidified, and centrifuged. Bottom panel: a TGFα-PE40 enriched sample.

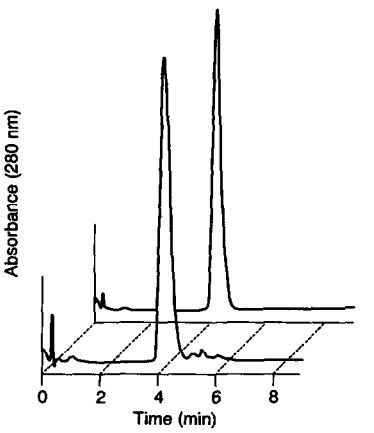

Fig. 2. Reversed-phase chromatograms of samples enriched in TGFα-PE40 or the $M_r \approx 40\,000$ protein. Approximately 7 μg of each protein were injected onto a HY-TACH non-porous C_{18} column and chromatographed as described under Experimental. The top chromatogram respresents TGFα-PE40 and the bottom chromatogram represents the $M_r \approx 40\,000$ protein.

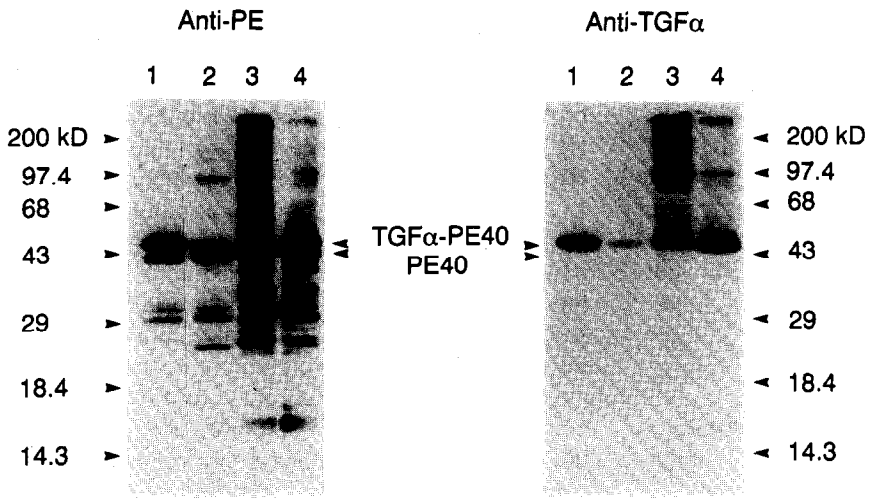

Fig. 3. Western blot analysis of TGFα-PE40 in-process samples. Proteins transferred to nitrocellulose paper were probed with either anti-PE or anti-TGFα antisera. The lanes for each blot correspond to: (1) a TGFα-PE40 enriched sample, (2) a PE40 enriched sample, (3) a cell lysate from *E. coli* that produced TGFα-PE40, and (4) a partially purified TGFα-PE40 sample. Lanes 1, 3, and 4 within each blot contained the same amount of TGFα-PE40 while lane 2 contained an equivalent amount of PE40. The gels were overloaded to highlight TGFα-PE40 aggregates and degradates in the samples. The migration of protein markers is indicated for each blot with masses expressed in kilodaltons (kD).

western blotting using antisera recognizing the PE40 portion of the chimeric protein showed that both proteins reacted with the antisera (Fig. 3). When western blots of these samples used anti-TGFα antisera, the sample enriched with TGFα-PE40 was recognized, but the $M_r \approx 40\,000$ protein was unreactive (Fig. 3). The absence of reactivity to anti-TGFα indicated that the $M_r \approx 40\,000$ protein did not contain TGFα epitopes. This finding, along with the approximate molecular size of the $M_r \approx 40\,000$ protein, prompted us to designate this impurity as PE40, a protein similar to that described by Chaudhary *et al.* [3]. It is not known whether the PE40 found in TGFα-PE40 samples was the result of proteolytic degradation or was synthesized *de novo* from an internal start signal in the messenger RNA.

The identical reversed-phase chromatographic behavior of TGFα-PE40 and PE40 might intially be considered surprising since the two proteins differ by approximately 50 amino acids. This suggests, however, that the 50 amino acid TGFα domain plays a minor role, if any, in the hydrophobic interaction between the protein and the stationary phase. Therefore, the PE40 portion of TGFα-PE40 must contribute significantly to the hydrophobic contact area. Since there are sequences within PE40 that are implicated in the proteins' ability to cross hydrophobic cell membranes [11], it is likely that these hydrophobic membrane-associating sequences are the principal means by which TGFα-PE40 and PE40 interact with the hydrocarbonaceous stationary phase. This interaction is so strong that any contribution from the TGFα domain is negligible.

Numerous attempts were made to resolve TGFα-PE40 from PE40 by reversed-phase HPLC using a number of different columns under a variety of mobile phase conditions. The types of columns tried included different silica-based columns with varying length alkyl side chains as well as wide-pore polystyrene–divinylbenzene polymeric resins. Various mobile phases at pH 2, 4.5, 7, 8.5, and 10 were evaluated, as were a number of organic modifiers. In all cases, TGFα-PE40 could not be fully resolved from PE40. Therefore, an alternative approach was sought to resolve the proteins. A high-resolution reversed-phase-based separation, however, was still desirable due to the need to monitor fermentation and crude samples. Examining the amino acid sequence of TGFα-PE40 revealed that all six of the

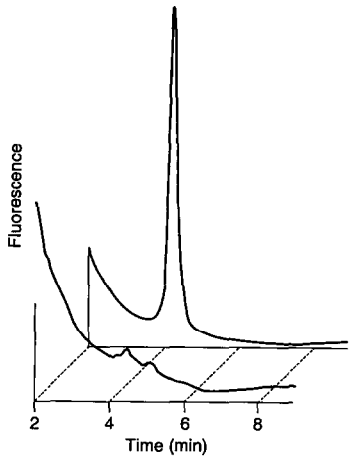

Fig. 4. Reversed-phase chromatograms of TGFα-PE40 and PE40 enriched samples labeled with mBBr. Approximately 7 μg of each protein were injected onto a HY-TACH non-porous C_{18} column and chromatographed as described under Experimental. The top chromatogram represents TGFα-PE40 and the bottom chromatogram represents PE40.

protein's cysteine residues were located in the TGFα domain [12]. This difference in the primary structure of the proteins was exploited by using the sulhydryl-specific fluorescent label mBBr. When samples highly enriched in TGFα-PE40 or PE40 were labeled with mBBr and then chromatographed by reversed-phase HPLC, mBBr-labeled TGFα-PE40 was detected by fluorescence whereas PE40 was not (Fig. 4). The small peak seen in the mBBr-treated PE40 chromatogram at 4.43 min indicates that the sample contained approximately 1% TGFα-PE40, which agrees with the anti-TGFα western blot analysis of that sample in Fig. 3, lane 2. Western blot analysis also revealed other low-molecular-mass fragments besides PE40 in the sample. Nevertheless, even if those fragments co-eluted with TGFα-PE40 by reversed-phase HPLC, the TGFα-PE40 would still be selectively identified since those protein fragments lacked the TGFα domain and therefore would not be labeled with mBBr.

Labeling TGFα-PE40 with mBBr in crude in-process samples, such as *E. coli* cell lysates, proved to be a powerful advantage of this technique. For these samples, labeling was rapid and showed the same dependence on mBBr concentration as it did for purified TGFα-PE40. Reversed-phase chromatograms of mBBr-labeled in-process samples are

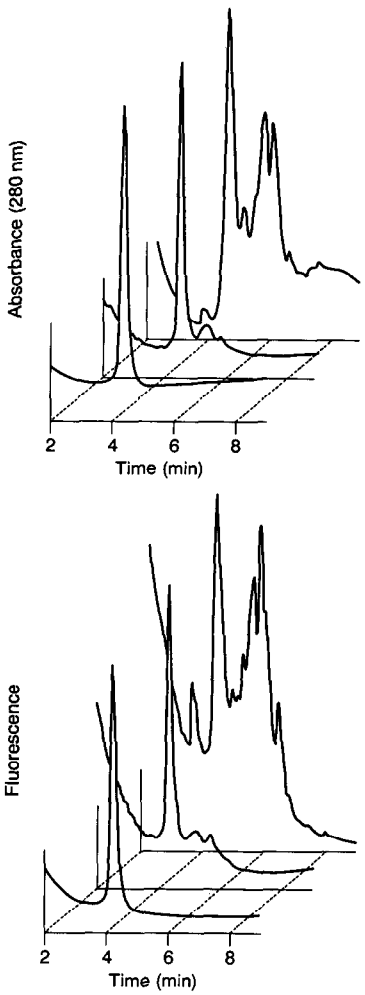

Fig. 5. Reversed-phase chromatograms of mBBr-labeled in-process samples. mBBr-labeled samples were chromatographed as described under Experimental. Column effluents were monitored for absorbance at 280 nm (upper panel) or fluorescence (lower panel). In each set of chromatograms the tracings are ordered from top to bottom as follows: a cell lysate from *E. coli* that produced TGFα-PE40, a partially purified sample of TGFα-PE40, and a TGFα-PE40 enriched sample.

shown in Fig. 5. In general, the profiles were similar whether the column effluent was monitored for absorbance at 280 nm (upper panel) or for fluorescence (lower panel). Control experiments demonstrated that the TGFα-PE40 contained in crude samples was completely labeled. The amount of TGFα-PE40 in each of these in-process samples was

TABLE I

IN-PROCESS QUANTITATION OF TGFα-PE40 BY ABSORBANCE AND FLUORESCENCE OF mBBr-LABELED SAMPLES

mBBr-labeled samples were chromatographed as described under Experimental and monitored simultaneously for absorbance at 280 nm and fluorescence. Standard curves were generated from the peak heights of known amounts of mBBr-labeled TGFα-PE40. The amount of TGFα-PE40 in unknown samples was interpolated from the two standard curves.

TGFα-PE40 Sample	Absorbance at 280 nm (mg/ml)	Fluorescence detection (mg/ml)
E. coli cell lysate	1.9	1.2
Partially purified sample	1.3	1.0
TGFα-PE40 enriched sample	1.6	1.6

interpolated from standard curves generated by chromatographing purified mBBr-labeled TGFα-PE40 and monitoring fluorescence and absorbance at 280 nm simultaneously. The TGFα-PE40 content in the cell lysate and the partially purified sample was lower by fluorescence detection than by absorbance at 280 nm (Table I). This was ture for many additional in-process samples. This suggested that proteins other than TGFα-PE40 were present in the TGFα-PE40 peak. A large component of this additional protein was PE40 (see Fig. 3). Therefore, quantitation without fluorescence detection would yield erroneous results leading to an overestimation of product levels. In the case of a highly enriched sample of TGFα-PE40, the assay showed that the TGFα-PE40 content was identical regardless of whether absorbance at 280 nm or fluorescence detection was used as the standard curve (Table I). Analysis of this same sample by Coomassie Blue staining of SDS-polyacrylamide gels verified the absence of PE40. These results demonstrate that mBBr-labeled samples provided the selectivity necessary to discriminate between TGFα-PE40 and PE40 allowing more accurate in-process quantitation of TGFα-PE40.

Although the reversed-phase separation originally provided the necessary efficiency to resolve TGFα-PE40 from host-cell proteins, it was possible that upon labeling the samples with mBBr the chromatographic behavior of other proteins could be affected and as a consequence, interfere with the TGFα-PE40 and PE40 peak. These labeled impurities would contribute to the TGFα-PE40 peak height for both UV absorbance and fluorescence detection leading to an overestimation of TGFα-PE40 levels by either detection method. Control experiments demonstrated that the UV absorbance chromatograms of mBBr-labeled in-process samples and unlabeled samples were not significantly different suggesting that the fluorescent tag did not alter the retention of labeled host-cell proteins.

The approach for selectively labeling TGFα-PE40 with mBBr for in-process monitoring and quantitation may be extended to other proteins that are difficult to resolve from impurities. Although the impurity described here may be unique, *i.e.* PE40 lacked sulfhydryls that were found in TGFα-PE40, there are a variety of other fluorescent compounds that are reactive with amine and carboxylic acid side chains in proteins [13,14]. A higher content of any one of these amino acid side chains in a protein compared to an impurity may provide enough discrimination to differentiate the two proteins. Fluorescent reagents that react with carbohydrates also exist [13]. This provides an additional avenue of analysis for recombinant proteins derived from mammalian cells. Whether any of these other fluorescent reagents are able to react quickly with their target protein under mild conditions in crude in-process samples remains to be determined. Finally, the absence of appropriate amino acid residues in recombinant proteins does not preclude the method described here. Specific amino acids can be engineered into proteins so that a fluorescent label may be attached to facilitate in-process monitoring and quantitation.

ACKNOWLEDGEMENTS

We gratefully acknowledge the support and en-

couragement of Drs. Hillel Cohen, Robert Sitrin and David Wonnacott. We are also grateful to C. Meacham Harrell and Anthony Paiva for assistance with polyacrylamide gel electrophoresis analysis and to Dr. Edgar Scattergood for his assistance throughout the course of this work.

REFERENCES

1 G. M. Edwards, D. Defeo-Jones, J. Y. Tai, G. A. Vuocolo, D. R. Patrick, D. C. Heimbrook and A. Oliff, *Mol. Cell. Biol.*, 9 (1989) 2860.
2 D. C. Heimbrook, S. M. Stirdivant, J. D. Ahern, N. L. Balishin, D. R. Patrick, G. M. Edwards, D. Defeo-Jones, D. J. FitzGerald, I. Pastan and A. Oliff, *Proc. Natl. Acad. Sci. U.S.A.*, 87 (1990) 4697.
3 V. K. Chaudhary, Y-H. Xu, D. FitzGerald and I. Pastan, *Proc. Natl. Acad. Sci. U.S.A.*, 85 (1988) 2939.
4 N. S. Kosower and E. M. Kosower, *Methods. Enzymol.*, 143 (1987) 76.
5 U. K. Laemmli, *Nature (London)*, 227 (1970) 680.
6 H. Towbin, T. Staehelin and J. Gordon, *Proc. Natl. Acad. Sci. U.S.A.*, 76 (1979) 4350.
7 R. K. Tweten and R. J. Collier, *J. Bateriol.*, 156 (1983) 680.
8 M. S. Blake, K. H. Johnston, G. J. Russell-Jones and E. C. Gotschlich, *Anal. Biochem.*, 136 (1984) 175.
9 K-J. Pluzek and J. Ramlau, in O. J. Bjerrum and N. H. H. Heegaard (Editors), *CRC Handbook of Immunoblotting of Proteins*, Vol. 1, CRC Press, Boca Raton, FL, 1988, p. 177.
10 J. Frenz, W. S. Hancock, W. J. Henzel and Cs. Horváth, in K. Gooding and F. Regnier (Editors), *HPLC of Biological Macromolecules, Methods and Applications*, Vol. 51, Marcel Dekker, New York, 1990, p. 145.
11 J. Hwang, D. J. FitzGerald, S. Adhya and I. Pastan, *Cell*, 48 (1987) 129.
12 D. Defeo-Jones, J. Y. Tai, R. J. Wegrzyn, G. A. Vuocolo, A. E. Baker, L. S. Payne, V. M. Garsky, A. Oliff and M. W. Riemen, *Mol. Cell. Biol.*, 8 (1988) 2999.
13 Y. Ohkura and H. Notha, *Adv. Chromatogr. (NY)*, 29 (1989) 221.
14 R. F. Chen and C. H. Scott, *Anal. Lett.*, 18 (A4) (1985) 393.