Structural characterisation of native and recombinant forms of the neurotrophic cytokine MK

Louis Fabri and Hiroshi Maruta

Melbourne Tumour Biology Branch, Ludwig Institute for Cancer Research (Melbourne Branch),
P.O. Royal Melbourne Hospital, Melbourne, Victoria 3050 (Australia)

Hisako Muramatsu and Takashi Muramatsu

Department of Biochemistry, Kagoshima University Medical School, 8-35-1 Sakuragaoka, Kagoshima 890 (Japan)

Richard J. Simpson

Joint Protein Structure Laboratory, Ludwig Institute for Cancer Research and The Walter and Eliza Hall Institute,
P.O. Royal Melbourne Hospital, Melbourne, Victoria 3050 (Australia)

Antony W. Burgess and Edouard C. Nice*

Melbourne Tumour Biology Branch, Ludwig Institute for Cancer Research (Melbourne Branch),
P.O. Royal Melbourne Hospital, Melbourne, Victoria 3050 (Australia)

ABSTRACT

The retinoic acid (RA)-inducible midkine (MK) gene encodes a heparin-binding protein which can induce neurite outgrowth in cultured mammalian embryonic brain cells. This cytokine shares 65% amino acid sequence identity with another RA-inducible cytokine, pleiotropin (PTN). Both proteins contain 10 conserved cysteine residues, all of which appear to be disulphide linked. MK and PTN are also rich in lysine and arginine residues rendering them susceptible to proteolysis during purification, and making large-scale preparation of these molecules inherently difficult.

Recombinant MK has been expressed as a fusion protein using a pGEX vector transfected into *E. coli*. To enable refolding of MK, the fusion protein was stored in solution at 4°C for 14 days in the presence of dithiothreitol (DTT). Thrombin cleavage of the fusion protein, post storage, typically generated 5 mg of MK per litre of bacterial pellet. To establish the structural integrity of the recombinant product, we have analysed the refolding kinetics and compared the disulphide bond assignment of recombinant MK with that of native MK and native PTN.

The synergistic use of micropreparative HPLC, to separate and recover in small eluant volumes enzymatically derived peptide fragments, with matrix assisted laser desorption mass spectrometry (MALD-MS) and N-terminal sequence analysis has allowed the unambiguous identification of the disulphide bonded fragments of native and recombinant MK. The disulphide bond assignment of MK is C_{12}–C_{36}, C_{20}–C_{45}, C_{27}–C_{49}, C_{59}–C_{91} and C_{69}–C_{101}, and is equivalent to that of PTN.

INTRODUCTION

Retinoic acid (RA) is a key molecule in the control of cell differentiation and development [1]. It has been shown to play an important role in morphogenesis during limb formation [2] and stimulates differentiation in epithelial and hemopoietic cells [1,3]. RA has also been proposed for use as a therapeutic agent for tumours

* Corresponding author.

capable of differentiation [1,4], but, due to its teratogenic effects, the use of RA is likely to be restricted. Midkine (MK), which was originally discovered from the analysis of mRNA from RA-inducible embryonal carcinoma cells [5], has been shown to promote the outgrowth of foetal brain neurones [6–8] and also promote both proliferation and neurite outgrowth of PC12 cells [8]. Native MK has also been reported to have mitogenic activity on NIH/3T3 fibroblasts [6]. These mitogenic effects were somewhat unexpected [4,9], and could be due to contamination by fibroblast growth factor, PTN or other as yet unidentified heparin binding cytokines which are present in preparations of native MK [4,9].

It is possible that MK, a product of RA induction, may well have similar therapeutic effects without the associated problems of teratogenicity. The production of large quantities of biologically active protein for such studies using recombinant techniques would therefore be advantageous.

We have described previously [4] a protocol for the production of large (mg) quantities of recombinant MK. Two difficulties arose during large-scale production of MK. Firstly, the high degree of lysine and arginine residues in MK make the protein susceptible to proteolytic degradation during the purification and secondly, the 10 cysteine residues have to be oxidised and arranged correctly to form the native structure. MK has been reported to form stable, non-native disulphide conformations during production in bacteria [4,9]. Incorrect folding of the MK is anticipated to reduce biological activity. However, current biological assay techniques for MK are limited and difficult to quantitate. Consequently it is difficult to assess the structural integrity using specific activity calculations.

We have therefore determined the disulphide bond assignment of recombinant MK and compared this assignment with that of both native MK and PTN [10]. The disulphide bond-containing fragments were generated by proteolytic digestion, purified by micropreparative HPLC and analysed by matrix assisted laser desorption mass spectrometry (MALD-MS) and N-terminal sequence analysis.

MALD-MS, originally described by Karas and

Hillenkamp [11], utilises a UV-absorbing matrix mixed with a sample to absorb energy applied by a pulsed UV laser. The sample is ionized and liberated into the gas phase and the mass analysis determined by time-of-flight mass spectrometry. MALD-MS has the advantageous features of being able to analyse mixtures of proteins [12,13], it is relatively insensitive to salts (a key feature of the technique first noted by Beavis and Chait [14]) and has small sample volume requirements making the technique ideally suited to the direct analysis of samples purified by micropreparative reversed-phase (RP) HPLC.

EXPERIMENTAL

Purification of native MK and PTN

The purification of native MK from murine L cells has been described elsewhere [6]. PTN was purified from bovine brain [10] using a modification of the method described by Bohlen et al. [15]. Briefly, the protein was purified by sequential chromatographic steps employing cation exchange on CM-Sepharose, affinity chromatography on heparin-Sepharose (Kabi Pharmacia, Uppsala, Sweden) and RP-HPLC on Brownlee RP-300 (100 × 4.6 mm I.D., Applied Biosystems, Foster City, CA, USA) [10]. Homogeneity of purified fractions was determined by sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS-PAGE), MALD-MS and N-terminal sequence analysis.

Purification of recombinant MK

Recombinant MK was purified as described previously by Maruta et al. [4]. Briefly, MK was expressed in Escherichia coli as a glutathione S-transferase fusion protein that was localised in the inclusion bodies of cellular extracts. The protein was recovered as an insoluble pellet which required 8 M guanidine hydrochloride (GuHCl) for solubilisation. The protein solution was then dialysed against thrombin cleavage buffer: 20 mM Tris–HCl pH 7.5, 2 mM $MgCl_2$, 1 mM dithiothreitol (DTT) (Sigma, St. Louis, MO, USA) and 1 mM $CaCl_2$. The sample was left for 2 weeks at 4°C prior to cleavage with thrombin. The resultant recombinant MK, which is insoluble at this stage, was solubilised in 8 M

GuHCl, 50 mM Tris–HCl pH 7.5 and subsequently purified by preparative RP-HPLC on a Vydac C_4 column (100 × 10 mm I.D.). This protocol typically generated 5 mg of purified MK per litre of bacterial pellet.

Time course of refolding of recombinant reduced MK

Purified recombinant MK (40 μg) was evaporated to dryness (Speedy-Vac concentrator, Savant Instruments, Farmingdale, NY, USA) and was redissolved in 6 M GuHCl, 1 mM EDTA, 0.25 M Tris–HCl pH 8.5 (100 μl) prior to reduction with DTT (10 μl, 5 mg/ml) for 1 h under nitrogen at ambient temperature. The reduced MK was desalted, and separated from excess low-molecular-mass reagents, using a Pharmacia Fast Desalting PC 3.2/10 column in a micropreparative HPLC fitted with an on-line conductivity meter (Pharmacia SMART system, Pharmacia LKB Biotechnology, Uppsala, Sweden). The eluant was 10 mM HCl and the flow-rate was 190 μl/min. The sample was recovered manually, with an allowance made for the post-detector dead volume. Collection was ceased directly the eluent conductivity began to increase.

The sample (190 μl) was diluted to 1 ml with the following oxido-shuffling buffer system: 0.2 M Tris–HCl pH 8.5, 1 mM EDTA, 1 M GuHCl, 0.6 mM oxidised glutathione and 6 mM reduced glutathione. Refolding was allowed to proceed at room temperature, and at appropriate timed intervals the reaction was stopped and the resultant products analysed by direct injection onto RP-HPLC.

RP-HPLC purification of proteins and peptides

Micropreparative RP-HPLC was performed using a Hewlett-Packard Model 1090A liquid chromatograph fitted with a Model 1040A diode-array detector and a Model 79994A computerised workstation (Hewlett-Packard, Waldbronn, Germany) as described elsewhere [16]. Reversed-phase supports used were Vydac C_4 (Separations Group, Hisperia, CA, USA), Brownlee RP-300 (100 × 4.6 mm I.D and 30 × 2.1 mm I.D., Applied Biosystems) and Hypersil-ODS (100 × 2.1 mm I.D., Hewlett-Packard). Samples

were recovered by manual collection into 1.5-ml screw-top polypropylene tubes (Sarstedt, Newton, NC, USA) and stored at 4°C prior to analysis. Chromatographic conditions were as described in the appropriate figure legends.

Enzymatic fragmentation of PTN and MK

Native PTN (40 μg), native MK (100 μg) and recombinant MK (40 μg) were digested with sequence-modified trypsin (Promega, Madison, WI, USA) in 1% (w/v) ammonium hydrogencarbonate containing 0.02% (v/v) Tween 20 (Pierce, Rockford, IL, USA) at an enzyme/substrate ratio of 1:20. The resultant peptides were purified by RP-HPLC as described above. Where peptides contained more than one disulphide bond due to the lack of internal tryptic cleavage sites, further digestion was performed using other enzymes. In the case of PTN a core fragment containing 6 of the 10 cysteines was digested further using either Asp-N endoprotease or chymotrypsin [10]. An MK core peptide containing 4 of the 10 cysteines was fragmented further using *Staphylococcus aureus* V8 protease (ICN ImmunoBiochemicals, Lisle, IL, USA) in 50 mM ammonium acetate pH 4.0 containing 0.02% (v/v) Tween 20 at an enzyme/substrate ratio of 1:10. Following RP-HPLC purification, peptide identity was determined by comparison of molecular mass of the peptides, derived by MALD-MS, with the calculated molecular mass determined from published amino acid sequence. The identity was further confirmed using N-terminal sequence analysis.

Determination of free sulphydryls using 5-I-AEDANS

Recombinant MK, which had been purified by RP-HPLC and dried using a Speedi-Vac concentrator, was reconstituted in 50 μl 6 M GuHCl in 0.2 M Tris–HCl buffer, pH 8.0, containing 1 mM EDTA. 5-N-[(Iodoacetamidoethyl)amino]-naphthalene-1-sulphonic acid (5-I-AEDANS) was then added (50 μl, 100 mM) and the reaction allowed to proceed for 5 min at room temperature in the dark. The reaction was stopped and the product desalted and separated from excess reagents by direct injection onto a Pharmacia Fast Desalting PC 3.2/10 column

fitted in a Pharmacia SMART system as described above. The desalted fraction was analysed by RP-HPLC whilst monitoring the absorbance at 215, 280 and 340 nm. Samples containing sulphydryls exhibit absorbance at all three wavelengths.

MALD-MS

Mass spectra were obtained using a matrix assisted laser desorption mass spectrometer, (Lasermat, Finnigan-MAT, San Jose, CA, USA) previously calibrated with synthetic peptides of known molecular mass ranging from 960 to 5200.

An aliquot (1 μl) of the sample from RP-HPLC was dispensed into a 1.5-ml test tube and mixed with 1 μl of suitable matrix: 3,5-dimethoxy-4 hydroxy cinnamic acid (sinapinic acid, 10 mg/ml in 70% (v/v) aqueous acetonitrile, Finnigan-MAT) or α-cyano-4 hydroxy cinnamic acid (10 mg/ml in 30% (v/v) aqueous acetonitrile, Aldrich, Milwaukee, WI, USA). The mixture was then applied to a target probe and allowed to dry at room temperature before insertion through a vacuum lock into the mass spectrometer. Laser power was arbitrarily set and fired until a spectrum was apparent. Laser power was then reduced to the minimal energy required to ionise and liberate the protein/matrix into gas phase and obtain a discernible mass spectrum. Accurate mass spectra were achieved by firing the laser until a discrete constant signal of minimal mass was produced. Peak detection and characterisation were performed using the instrumentation data evaluation programs.

SDS-PAGE

Samples were prepared for SDS-PAGE analysis as previously described [17] and separated on a PhastSystem (Pharmacia LKB, Sweden) using 10–15% Phastgels. Proteins were visualised by silver staining [18] using the modification for semi-automated staining in the PhastSystem described below.

Step	Reagent	Time (min)	Temperature (°C)
1	Methanol–acetic acid–water (50:10:40, v/v/v)	5	50
2, 3, 4	Ethanol–acetic acid–water (10:5:85, v/v/v)	5	50
5	0.2% (w/v) Potassium dichromate and 0.2% (v/v) nitric acid in water	5	50
6	0.2% (w/v) Silver nitrate in water	25	40

At the termination of the automated program the gels were removed from the development chamber and the staining completed manually, in a glass beaker, using 3% (w/v) sodium carbonate, 0.05% (v/v) formaldehyde. Gentle agitation was maintained, whilst viewing the gel using a light box, until a suitable staining level had been achieved. Staining was arrested by placing the gels in acetic acid–glycerol–water (10:10:80, v/v/v) for 5 min at room temperature.

N-Terminal amino acid sequence analysis

Edman degradation of proteins or peptides was performed using an automated amino acid sequencer (Models 470A or 477A, Applied Biosystems) equipped with an on-line phenylthiohydantion (PTH) amino acid analyser (Model 120A) as described previously [19].

RESULTS AND DISCUSSION

The primary sequence of MK has been deduced for human [7,20], chicken [21] and mouse [22] from the corresponding cDNA. There is a high degree of homology between species (87, 65 and 64% homology between human and mouse, human and chicken or mouse and chicken MK, respectively). Human MK shares 65% homology with another cytokine, PTN, which has been isolated from bovine brain [10,15]. PTN and MK are both rich in lysine and arginine residues and have 10 conserved cysteine residues (Fig. 1). All the cysteine residues in PTN have been reported to form disulphide bonds [10,23]. We have recently determined the topology of the disulphide bonds in PTN [10] as described under Experimental. Disulphide bonds were located between $C_{15}–C_{44}$, $C_{23}–C_{53}$, $C_{30}–C_{57}$, $C_{67}–C_{99}$ and $C_{77}–C_{109}$.

The possibility of improper disulphide bond formation is a common problem during the production of bacterially synthesized proteins

```
bPTN    (1)   GKKEKPEKKVKK-SDCGEWQWSVCV
mMK     (1)   KKKEK----VKKGSECSEWTWGPCT

bPTN   (25)   PTSGDCGLGTREGTRTGAECKQTMK
mMK    (22)   PSSKDCGMGFREGT----CGAQTQ

bPTN   (50)   TQRCKIPCNWKKQFGAECKYQFQAW
mMK    (42)   RVHCKVPCNWKKEFGADCKYKFESW

bPTN   (75)   GECDLNTALKTRTGSLKRALHNADC
mMK    (67)   GACDGSTGTKARQGTLKKARYNAQC

bPTN  (100)   QKTVTISKPCGKLTKSKPQAESKKK
mMK    (92)   QETIRVTKPCTSKTKSKTKAKKGKG

bPTN  (125)   KKEGKKQEKMLD
mMK   (117)   KD
```

Fig. 1. Comparison of the deduced amino acid sequences of murine MK [22] and bovine PTN [10]. The conserved cysteine residues of PTN and MK are boxed in black. The lysine and arginine residues, which render the protein susceptible to proteolysis during purification, are highlighted in grey.

Fig. 2. N-Terminal sequence analysis of recombinant MK purified by RP-HPLC. An aliquot of recombinant MK (10 μg), which had been purified by preparative RP-HPLC, was rechromatographed using microbore RP-HPLC. The sample was recovered as indicated by the horizontal bar. The N-terminal sequence data obtained is shown (inset top). A 2.5-μl aliquot of the sample was analysed by SDS-PAGE (shown inset bottom) as described under Experimental. Chromatographic conditions: column, Brownlee RP-300 (30 × 2.1 mm I.D.); linear gradient of 0 to 100% solvent B over 60 min where solvent A was 0.15% (v/v) trifluoroacetic acid (TFA) in water and solvent B was acetonitrile–water (60:40, v/v), containing 0.125% (v/v) TFA. The flow-rate was 100 μl/min, and the temperature 45°C. kDa = kilodalton.

which are rich in cysteine [24], and heterogeneity of recombinant MK during large-scale production has been reported previously [4,7]. We have recently described [4] a protocol for the production of recombinant MK. In order to confirm the structural integrity of the recombinant product, an aliquot of recombinant MK (10 μg) was prepared for N-terminal sequence analysis using micropreparative RP-HPLC (Fig. 2). The sample chromatographed as a major species, with a characteristic elution time of 25.2 min, and appeared to be greater than 95% homogenous. A minor species which eluted as a shoulder on the trailing edge of the peak, was spectrally identical to the major component and probably represented trace levels of an incorrectly folded form.

The single N-terminal sequence (Fig. 2, top inset) confirmed the presence of the leader sequence (GS), arising from the thrombin specific cleavage sequence located between the C-terminus of the GST fusion protein and the mature form of murine MK. This was followed by the anticipated N-terminal sequence; KKKEKVKKGSE, confirming that the protein had not undergone proteolysis in the N-terminal region during production and purification. SDS-PAGE analysis (Fig. 2, bottom inset) indicated a single species migrating with an apparent molec-
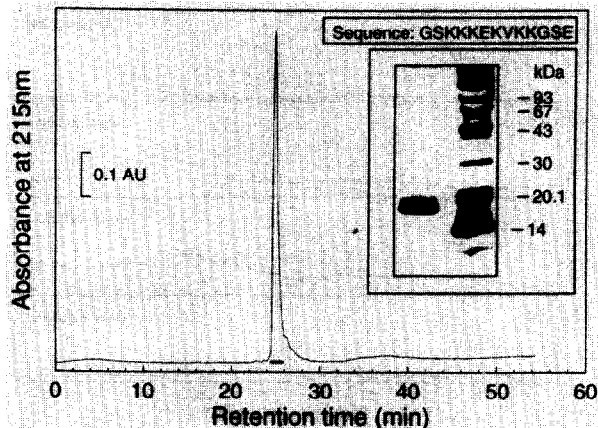
ular mass of 19 000. The difference in apparent molecular mass by SDS-PAGE and the calculated molecular mass of 13 226 from the amino acid sequence has been attributed to the highly basic nature of MK [22]. MALD-MS analysis of the recombinant product indicated a mean molecular mass of 13 200, supporting the authenticity of the product. Amino acid analysis (data not shown) was consistent with the integrity of the purified recombinant MK.

In order to optimise the folding of MK during the production protocol, an analysis of the kinetics of refolding was carried out. Recombinant MK was reduced with DTT and the product characterised by RP-HPLC (Fig. 3A). Reduced MK eluted at 30.8 min and was more hydrophobic on a C_8 reversed-phase support than the oxidised MK which eluted at 25.2 min under the same chromatographic conditions (Fig. 2). After 5 min in oxido-shuffling buffers (as described under Experimental) the level of the reduced material was considerably lower (note the 5-fold
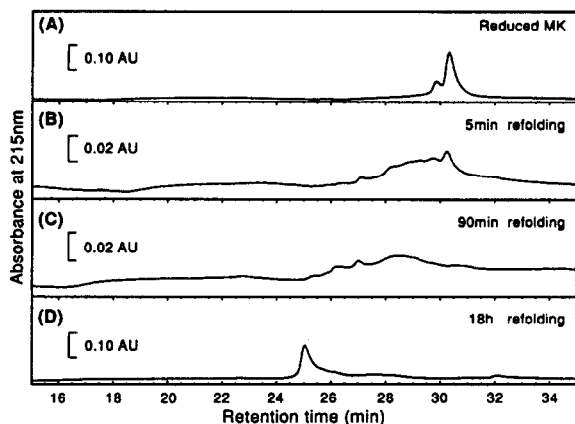
Fig. 3. Analysis of the refolding of reduced recombinant MK. Reduced recombinant MK was refolded as described under Experimental. At time 0, before the addition of the oxido-shuffling buffers, an aliquot of the reduced MK was injected onto the RP-HPLC columns (A). After 5 min, 90 min and 18 h incubation in the presence of the oxido-shuffling buffers equivalent aliquots were analysed (B, C and D, respectively). In all cases the sample was injected directly onto the column. Chromatographic conditions were as described in the legend to Fig. 2.

increase in sensitivity between Fig. 3A and B) and a number of less hydrophobic species were now present. This is thought to be due to "masking" of hydrophobic amino acids during the folding process. These species elute earlier since these residues are no longer available for hydrophobic interaction with the support material [25]. No species chromatographing with the characteristic retention time of 25.2 min were evident.

After 90 min (Fig. 3C) little fully reduced material remained, and there was a continued shift towards less hydrophobic intermediate forms of MK. After 18 h incubation (Fig. 3D) a major species eluting at 25 min was again apparent. This material was chromatographically indistinguishable from the recombinant MK produced by the 14-day aerobic folding method [4]. The overall yield of refolded material using the oxido-shuffling system was approximately 80% as calculated by the area under the curves (equivalent aliquots were injected in Fig. 3A, B, C and D). However, the 14-day aerobic folding protocol was preferred for the large-scale production of the recombinant MK because of reduced costs associated with the refolding re-

agents. Yields of MK using this protocol were typically 5 mg per litre of bacterial culture.

Biological assay of the recombinant MK [4] indicated that the product induced neurite outgrowth of central neurones from E18 mouse spinal cord and caused clumping of neurofilaments in E10 neuroepithelial cells (data not shown), features characteristic of the neurotrophic cytokines [4]. However, these assays are not readily quantifiable and hence accurate specific activity data is difficult to obtain. No mitogenic activity of the recombinant product on Balb/3T3 cells was detected, suggesting that the mitogenic activity observed in MK produced from biological sources [9] was probably due to contamination with low levels of other mitogenic species.

The above data indicates that the recombinant MK was homogenous with respect to composition, but does not confirm authenticity of the tertiary structure. The disulphide arrangement of the recombinant MK was analysed, and compared to native MK derived from mouse L cells [6] and native PTN from bovine brain [10].

Disulphide assignment of native MK

Inspection of the MK sequence (Fig. 1) indicated that the lysine and arginine distribution was suited to cleavage by trypsin. Given that the tryptic digest was complete, cysteine containing
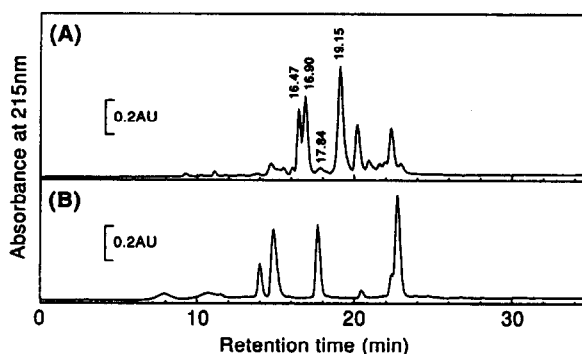

Fig. 4. Tryptic map of recombinant MK. Recombinant MK (20 μg) was digested with trypsin (1 μg) as described under Experimental. Chromatographic conditions were as described in Fig. 2. (A) RP-HPLC analysis of an aliquot (50%) of the tryptic digest of recombinant MK. (B) An equal volume of the tryptic digest was reduced with DTT before RP-HPLC as described under Experimental.

peptides comprising residues 9–25, 26–32, 33–42, 43–46, 47–52, 54–60, 63–76, 87–96 and 97–104 would be generated. All these tryptic peptides would contain a single disulphide bond, except peptide 9–25 where no cleavage site is present between the first and second cysteines (C_{12} and C_{20}). Peptide 9–25 could potentially be cleaved further using V8 protease at the glutamic acid residue at position 13, fragmenting the "core" peptide into the individual cysteine bonded fragments.

Recombinant MK (10 μg) was digested with trypsin (as described under Experimental) and the peptides analysed by microbore RP-HPLC (Fig. 4A). Five major peptides eluting at 16.47, 16.90, 17.84 and 19.15 min were detected. To determine which of these peptides contained disulphide bonds, an aliquot of the tryptic digest was reduced with DTT and separated using the same chromatographic conditions (Fig. 4B). All the major peak retention times were significantly altered on exposure to DTT demonstrating that these fractions most probably contained disulphide bonded peptides.

Second-order derivative spectra of the peptides, acquired during the chromatographic separation using a diode array detector, helped to identify the major peaks (Fig. 5). It is possible to identify tryptophan and tyrosine containing peptides by their characteristic minima at 290 ± 2 nm and 280 ± 2 nm respectively [26]. The peak eluting at 16.47 min had a minimum at 282 nm (Fig. 5A), corresponding to the presence of a tyrosine residue. Inspection of the amino acid sequence of MK (Fig. 1) suggests that this peptide fragment contained residues 87–96 as one component of the dipeptide since this is the only expected tryptic peptide to contain a tyrosine residue. The peaks eluting at 16.90, 17.84 and 19.15 minutes (Fig. 5B, C and D) all had second derivative minima at 290 nm indicating the presence of tryptophans. Inspection of the MK sequence (Fig. 1) predicts that the tryptophan residues would all be associated with disulphide-bonded tryptic dipeptides.

Native MK (100 μg) from mouse L cells [6] was rechromatographed using RP-HPLC to ensure homogeneity (Fig. 6A, inset). The sample was recovered manually (peak volume 140 μl) and dried at room temperature using a Speedi-Vac concentrator. After redissolving the sample in the trypsin digest buffer (see Experimental) and digesting overnight at 37°C with trypsin, the resultant peptides were separated by RP-HPLC (Fig. 6A). Peaks I/II, III and IV were collected manually and rechromatographed on Hypersil ODS (100 × 2.1 mm I.D.) (Fig. 6B). This procedure resulted in further purification, and concentrated the samples sufficiently for direct application to MALD-MS and N-terminal sequence analysis. Peaks IVb and the peak eluting at 20.2 min were analysed directly by MALD-MS. The peak eluting at 20.2 min was identified as residual undigested MK. The identity of peak IVb is discussed later.

The resultant mass spectra (Fig. 6C) enabled the rapid identification of the disulphide-bonded fragments. Peaks I, II and III corresponded to the peptide 54–60 linked by a disulphide bond (SS) to peptide 87–96, peptide 63–76 (SS) 97–104 and peptide 26–32 (SS) 47–52, respectively (see Table I). These peptides contained disulphide links between cysteines C_{59}–C_{91}, C_{69}–C_{101} and C_{27}–C_{49}, respectively. The identity of these peptides was confirmed by N-terminal sequence analysis.
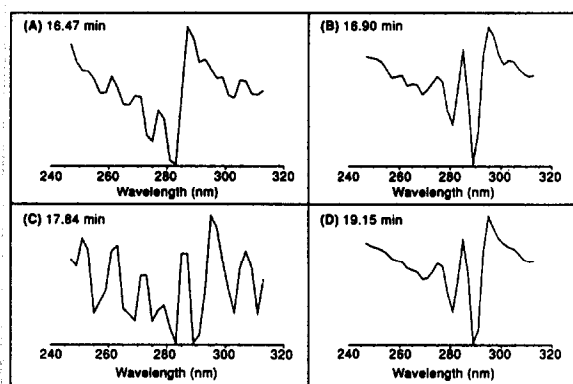
The mass spectral data for these peptides were

Fig. 5. Second order derivative spectroscopy of tryptic fragments of recombinant MK. The second-order derivative spectra of the major tryptic peptides of recombinant MK (Fig. 4) were obtained using an on-line diode array detector [26]. Tyrosine containing peptides show a characteristic minima of 280 ± 2 nm (see 16.47-min peak), whilst tryptophan containing peptides have minima at 290 ± 2 nm (see 16.90-, 17.84- and 19.15-min peaks).
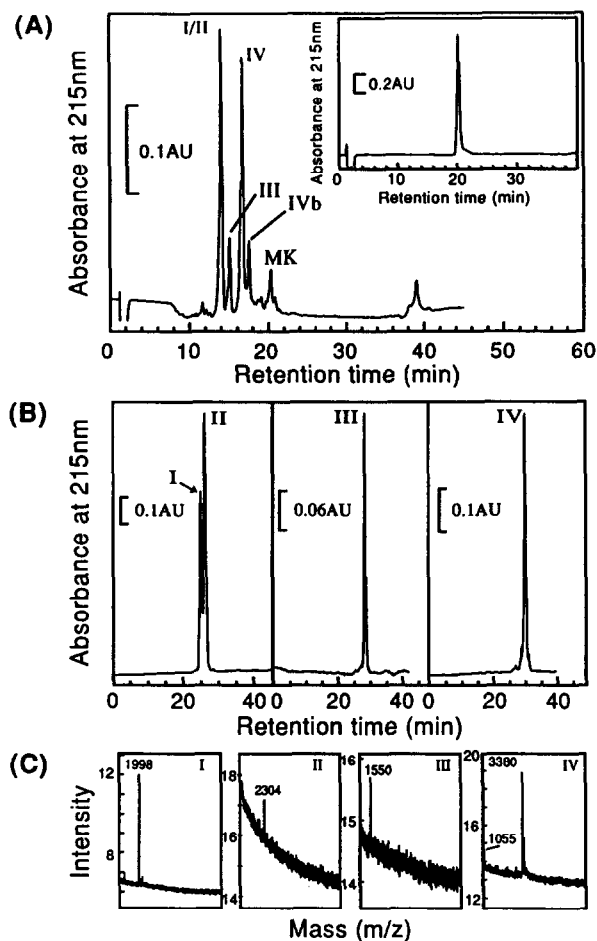
Fig. 6. Chromatographic purification and MALD-MS characterisation of tryptic peptides from native MK. (A) Native MK (100 μg) was digested with trypsin for 16 h at 37°C and the resultant peptides purified using a Brownlee RP-300 (100 × 4.6 mm I.D.) column. The chromatographic conditions were as described in Fig. 2 with the exception that the flow-rate was 1 ml/min. Inset: native MK (100 μg) was chromatographically purified prior to tryptic digestion using the same chromatographic conditions. (B) Rechromatography of tryptic peptides I–IV (Fig. 6A) using a Hypersil ODS (100 × 2.1 mm I.D.) column. Chromatographic conditions were as for Fig. 2. Purified peptides were characterised by MALD-MS and N-terminal sequence analysis. (C) Characterisation of purified tryptic peptides [I–IV in (B)] using MALD-MS with sinapinic acid as the matrix.

acquired using sinapinic acid as the matrix. This matrix frequently demonstrated poor signal-to-noise characteristics and often required high levels of laser energy to achieve adequate ionisa-

tion. This resulted in slightly inaccurate molecular mass analysis (with observed values characteristically higher than the theoretical molecular mass). The deviation from the theoretical mass values typically ranged from 0.1 to 0.2%. Peak III, peptide 26–32 (SS) 47–52, which gave the anticipated N-terminal sequence data, is believed to contain a methionine sulphoxide residue which accounts for the observed molecular mass difference of 16 mass units from that expected from the reported sequence.

An additional advantage of MALD-MS is the ability to readily identify and characterise impurities which would otherwise confound interpretation of N-terminal sequence data. The MALD-MS for peak II (Fig. 6C), when run at higher laser-energy levels, indicated the presence of residual levels of peak I (Fig. 6A). Unambiguous assignment of the N-terminal sequence data for this peak, where 2 major and 2 minor sequences were present, was therefore possible. The ability to identify specific contaminants in a sample, and account for them when assigning N-terminal sequence data, further enhances the use of MALD-MS for peptide analysis. One caveat should be noted: variable ion suppression may occur with MALD-MS [27] which can prevent the relative quantitation of impurities from the observed levels in the mass spectra.

In most cases the concentration of the peaks obtained from micropreparative HPLC permitted direct application of the sample (1 μl) to the target disk of the MALD-MS. This represents approximately 1% of the eluent peak volume (typically 100 μl or less). In some cases, where minor components of the mixture were analysed by MALD-MS, prior concentration of the sample using a Speedi-Vac concentrator was required.

Peak IV was identified as a "core peptide" containing four cysteine residues and which comprised of three peptides: residues 9–25, 33–42 and 43–46 (Fig. 1). Peak IVb (Fig. 6A) differed from peak IV by the presence of the single lysine at the N terminus of peptide 8–25.

The core peptide (IV) was digested using V8 protease and the peptides separated by RP-HPLC using a Hypersil ODS column (100 × 2.1 mm I.D.) (Fig. 7A). Three chromatographic

TABLE I

IDENTIFICATION OF THE PEPTIDES FROM THE TRYPTIC DIGEST OF NATIVE MK

Peak No. (Fig. 6)	Deduced sequence[a]	Theoretical molecular mass	Observed molecular mass[b]
I	EFGADC$_{59}$K YNAQC$_{91}$ETIR	1993	1997
II	FESWGAC$_{69}$DGSTGTK VTKPC$_{101}$TSK	2307	2304
III	DC$_{27}$GMGFR VPC$_{49}$NWK	1530	1549[c]
IV	GSEC$_{12}$SEWTWGPC$_{20}$TPSSK EGTC$_{36}$GAQTQR VHC$_{45}$K	3375	3379
IVb	KGSEC$_{12}$SEWTWGPC$_{20}$TPSSK EGTC$_{36}$GAQTQR VHC$_{45}$K	3503	3707

[a] Amino acid sequence identity was deduced from mass spectral analysis with reference to the mass calculated from the amino acid sequence (Fig. 1). The deduced sequence was confirmed by N-terminal amino acid sequence analysis.

[b] Observed molecular masses obtained from MALD-MS are expressed as the single ionised form (M^+).

[c] The difference between the theoretical and observed molecular mass is attributed to the presence of a methionine sulphoxide (residue 29).
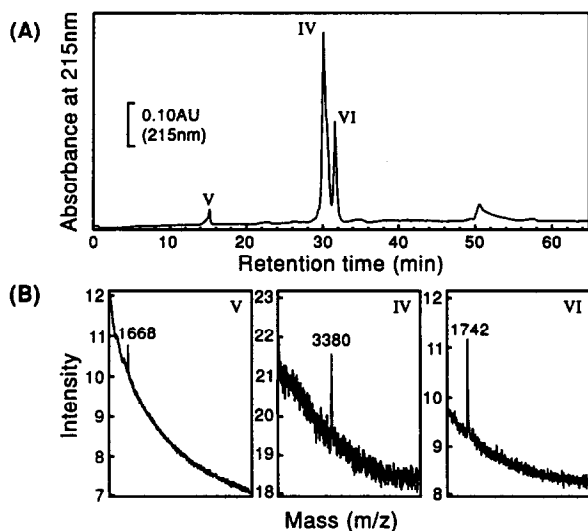

Fig. 7. S. aureus V8 protease digestion of the tryptic peptide core of native MK. (A) The MK tryptic core peptide (Fig. 6: peak IV) was digested with S. aureus V8 protease as described under Experimental, and the resultant peptides fractionated using a Hypersil ODS (100 × 2.1 mm I.D.) column. Chromatographic conditions were as described in Fig. 2. (B) The peptide fractions IV–VI were analysed by MALD-MS using sinapinic acid as the matrix.

peaks were detected: peak IV (the undigested core peptide) and peaks V and VI.

Peaks V and VI were characterised by MALD-MS (Fig. 6B). The disulphide bond assignment corresponding to the observed masses is shown in Table II. Peaks V and VI corresponded to residues 9–14 (SS) 33–42 and 15–26 (SS) 43–46, respectively. The mass accuracy was 0.2–0.4%.

The relative proportion of peak IV was high due to incomplete digestion by the V8 protease. The digestion was performed at pH 4.0 since when cleavage was performed at pH 8.5 disulphide interchange occurred. In particular a peptide containing a disulphide bond between C_{12} and C_{20} was generated. This peptide could not have been present originally in the core because it would have chromatographed as a discrete component in the original tryptic digest. Consequently, even though the kinetics were slower and the reaction incomplete, to ensure the authenticity of the disulphide bonds the V8 digestion was performed at pH 4.0.

The data obtained using the above protocol resulted in a unique disulphide assignment for

TABLE II

IDENTIFICATION OF THE PEPTIDES FROM THE V8 DIGEST OF THE TRYPTIC CORE OF NATIVE MK

Peak No. (Fig. 7)	Deduced sequence[a]	Theoretical molecular mass	Observed molecular mass[b]
V	GSEC$_{12}$SE EGTC$_{36}$GAQTQR	1660	1667
VI	WTWGPC$_{20}$GTPSSK VHC$_{45}$K	1736	1741

[a] Amino acid sequence identity was deduced from mass spectral analysis with reference to the mass calculated from the amino acid sequence (Fig. 1). The deduced sequence was confirmed by N-terminal amino acid sequence analysis.
[b] Observed molecular masses obtained from MALD-MS are expressed as the single ionised form (M$^+$).

native MK corresponding to: C$_{12}$–C$_{36}$, C$_{20}$–C$_{45}$, C$_{27}$–C$_{49}$, C$_{59}$–C$_{91}$ and C$_{69}$–C$_{101}$ which is equivalent to that reported for PTN [10].

Disulphide assignment of recombinant MK

An aliquot of recombinant MK (40 μg) was digested with trypsin as described under Experimental, and the resultant peptides VII/VIII,

IX and X separated by RP-HPLC on an RP-300 (100 × 4.6 mm I.D.) column (Fig. 8A) followed by micropreparative HPLC, using Hypersil ODS (100 × 2.1 mm I.D.) (Fig. 8B). Each peptide was analysed by MALD-MS (Fig. 8C). The deduced sequences of these peptides are given in Table III. The mass spectra were obtained using α-cyano-4-hydroxy cinnamic acid as the matrix

TABLE III

IDENTIFICATION OF THE TRYPTIC PEPTIDES OF RECOMBINANT MK

Peak No. (Fig. 8)	Deduced sequence[a]	Theoretical molecular mass	Observed molecular mass[b]
VII	EFGADC$_{59}$K YNAQC$_{91}$ETIR	1993	1992
VIII	EFGADC$_{59}$K YNAQC$_{91}$ETIR	1993	1993
	FESWGAC$_{69}$DGSTGTK VTKPC$_{101}$TSK	2307	2309
IX	DC$_{27}$GMGFR VPC$_{49}$NWK	1530	1545[c]
X	GSEC$_{12}$SEWTWGPC$_{20}$GTPSSK EGTGAC$_{36}$QTQR VHC$_{45}$K	3375	3376

[a] Amino acid sequence identity was deduced from mass spectral analysis with reference to the mass calculated from the amino acid sequence (Fig. 1).
[b] Observed molecular masses obtained from MALD-MS are expressed as the single ionised form (M$^+$).
[c] The difference between the theoretical and observed molecular mass is attributed to the presence of a methionine sulphoxide (residue 29).
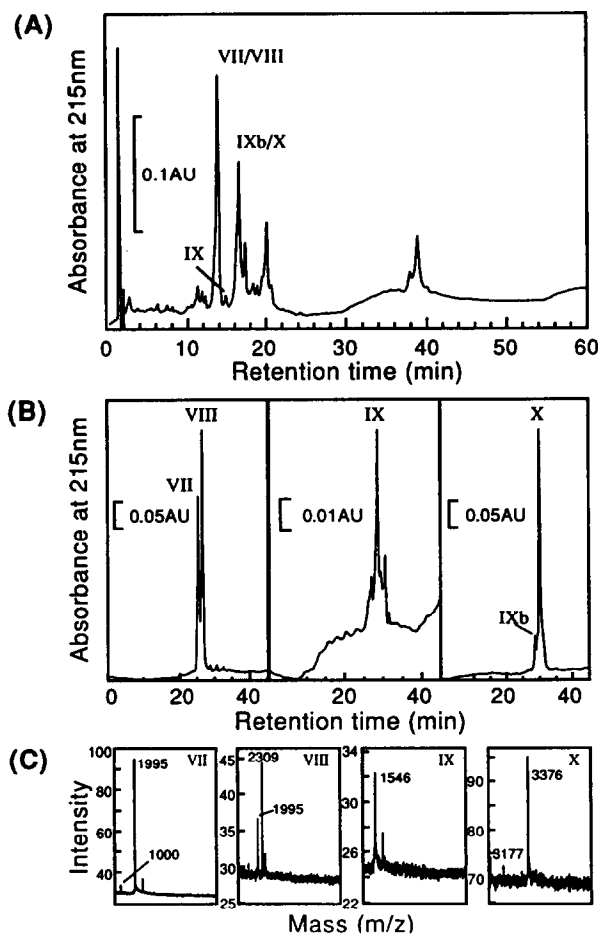
Fig. 8. RP-HPLC purification and MALD-MS characterisation of tryptic peptides of recombinant MK. (A) Recombinant MK (40 μg) was digested with trypsin for 16 h at 37°C. The resultant peptides, VII–X, were purified using a Brownlee RP-300 (100 × 4.6 mm I.D.) column as described in Fig. 6A. Peptides were recovered manually for rechromatography on a micropreparative column. (B) Peptides VII–X were rechromatographed using a Hypersil ODS column (100 × 2.1 mm I.D.) as described in Fig. 6B. (C) Peptides VII–X (Fig. 8B) were characterised using MALD-MS, using α-cyano-4-hydroxy cinnamic acid as the matrix.

[27,28] which gave more accurate mass assignments (better than 0.05%). The improvement in the mass accuracy is attributed to the lower laser energy required for ionisation with this matrix [29].

Peptide X (Fig. 8A), the core peptide, was fragmented further into the individual disulphide-bonded components using V8 protease at

pH 4.0. MALD-MS analysis of the V8 peptides, post purification on Hypersil ODS (data not shown), confirmed the presence of disulphide-bonded linkages corresponding to peptides 9–13 (SS) 33–42 and 14–25 (SS) 43–46. Taken together with the data from the tryptic digest the disulphide assignment of the recombinant MK was C_{12}–C_{36}, C_{20}–C_{45}, C_{27}–C_{49}, C_{59}–C_{91} and C_{69}–C_{101} and is identical to that of native MK.

Comparison of the tryptic maps for native MK (Fig. 6A) and recombinant MK (Fig. 8A) support the notion that the disulphide bond assignment of the recombinant and native proteins is very similar. With the exception of peak IX (Fig. 8A) and the presence of a non-retained peak at the solvent front (Fig. 8) the relative peak heights were identical. Initially we thought that the C_{27}–C_{49} bond might not be fully oxidised in the recombinant MK. However, no free sulphydryls could be detected using 5-I-AEDANS [30]. Either the sulphydryls were not accessible to the reagent due to stearic effects, or the recombinant MK was fully disulphide bonded and other structural modifications of the C_{27}–C_{49} dipeptide had taken place. If the disulphide bond had not been formed, additional chromatographic species would have been expected on the C_8 column. In particular the sequence VPCNWK should be sufficiently hydrophobic to bind to a C_8 support [31], and would have a characteristic second order derivative spectrum [26]. This peptide was not present in the tryptic digests, and further examination of the tryptic map was required to account for the low level of peak IX (40% of that anticipated). In particular we had deduced from the MALD-MS data that peak III in the native MK digest (Fig. 6) and peak IX in the recombinant digest (Fig. 8) contained an oxidised methionine residue (see Tables I and III, respectively). It therefore seemed possible that a peak corresponding to the non-oxidised form, or some other modification, might be present. Using MALD-MS to analyse peak IXb, a mass species with an identical mass to that of C_{27}–C_{49} was detected. The chromatographic shift of the peptide to a later retention time could be due to deamidation of asparagine (residue 50) to an aspartic acid. Such chromatographic behaviour on RP-HPLC has been noted

previously for insulin [32]. The relative change in mass for this modification (-1) is not sufficiently different from that of the oxidised material for unambiguous identification by MALD-MS. This peak accounts for a further 20% of the anticipated C_{27}–C_{49} peptide.

The conserved disulphide bond assignment of MK and PTN further suggests a close relationship between these neurotrophic cytokines. As well as showing a high degree of primary sequence homology, the similar disulphide assignments indicate that these proteins share similar tertiary structures. Far-UV circular dichroic (CD) spectra of PTN [10] showed a large proportion of β sheet (80%) and a small proportion of random coil. Far-UV CD spectra of MK also indicated high levels of β sheet and random coil (data not shown). Predictive studies, using the algorithm of Garnier et al. [33], indicated the secondary structure shown together with the disulphide bond assignments in Fig. 9.

The biological effects of MK and PTN, although distinct, may be regulated by similar binding sites within their respective receptor domains. It will be interesting to determine whether MK and PTN are analogous to epidermal growth factor and transforming growth factor α [34] or insulin and insulin-like growth factor(s) [35], which share sequence, functional and conformational homology. It is now important to determine the nature of the receptor interactions for these two closely related neurotrophic cytokines.
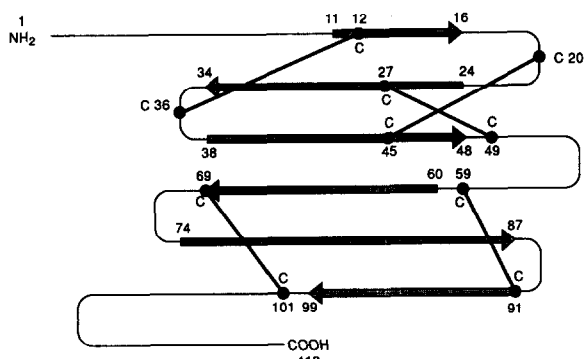
ACKNOWLEDGEMENTS

Fig. 9. Disulphide bond assignment and predicted secondary structure for MK. Prediction was made using the algorithm of Garnier et al. [33] parameterised to 60% β sheet as per the results of the analysis of the CD spectrum (data not shown). Probability cut-off parameters used were: α helix, 400; β sheet, 20; turn, 300; and random coil, 200. Predicted β sheets are shown as hatched arrows. Disulphide linkages are shown as bold lines.

REFERENCES

1 M.B. Sporn and A.B. Roberts, Cancer Res., 43 (1983) 3034.
2 C. Thaller and G. Eichele, Nature, 327 (1987) 625.
3 T.R. Breitman, S.J. Collins and B.R. Keene, Proc. Natl. Acad. Sci. U.S.A., 77 (1980) 2936.
4 H. Maruta, P. Bartlett, V. Nurcombe, M.S.A. Nur E Kamal, C. Chomienne, T. Muramatsu, H. Muramatsu, L. Fabri, E. Nice and A.W. Burgess, Growth Factors, 8 (1993) 119.
5 K. Kadomatsu, M. Tomomua and T. Muramatsu, Biochem. Biophys. Res. Commun., 151 (1988) 1312.
6 H. Muramatsu and T. Muramatsu, Biochem. Biophys. Res. Commun., 177 (1991) 652.
7 P.J. Kretschmer, J.L. Fairhurst, M.M. Decker, C.P. Chan, Y. Gluzman, P. Bohlen and I. Kovesdi, Growth Factors, 5 (1991) 99.
8 D. Raulais, O. Lagenete-Chevallier, C. Guettet, D. Duprez, Y. Courtois and M. Vigny, Biochem. Biophys. Res. Commun., 171 (1991) 603.
9 P. Bohlen and I. Kovesdi, Progr. Growth Factor Res., 3 (1991) 143.
10 L. Fabri, E.C. Nice, L.D. Ward, H. Maruta, A.W. Burgess and R.J. Simpson, Biochem. Int., 28 (1992) 1.
11 M. Karas and F. Hillenkamp, Anal. Chem., 60 (1988) 2299.
12 S.A. Carr, M.E. Hemling, M.F. Bean and G.D. Roberts, Anal. Chem., 63 (1991) 2802.
13 F. Hillenkamp, M. Karas, R.C. Beavis and T. Chait, Anal Chem., 63 (1991) 1193A.
14 R.C. Beavis and B.T. Chait, Anal. Chem., 62 (1990) 1836.
15 P. Bohlen, T. Muller, P. Gautschi-Sova, U. Albrecht, C.G. Rasool, M. Decker, A. Seddon, V. Fafeur, I. Kovesdi and P.J. Kretschmer, Growth Factors, 4 (1991) 97.

16 E.C. Nice, P. Dempsey, J. Layton, G. Morstyn, D. Cui, R. Simpson, L. Fabri and A.W. Burgess, *Growth Factors*, 3 (1990) 159.

17 E.C. Nice, L. Fabri, A. Hammacher, J. Holden, R.J. Simpson and A.W. Burgess, *J. Biol. Chem.*, 267 (1992) 1546.

18 C.R. Merril, D. Goldman, S.A. Sedman and M.H. Ebert, *Science*, 211 (1981) 1437.

19 G.S. Begg and R.J. Simpson, in T.E. Huglei (Editor), *Techniques in Protein Chemistry*, Academic Press, San Diego, CA, 1989, p. 79.

20 J. Tsutsui, K. Uehara, K. Kadomatsu and T. Muramatsu, *Biochem. Biophys. Res. Commun.*, 176 (1991) 792.

21 P. Urios, D. Duprez, J.P. Le Caer, Y. Courtis, M. Vigny and M. Laurent, *Biochem. Biophys. Res. Commun.*, 175 (1991) 617.

22 M. Tomomura, K. Kadomatsu, S. Matsubara and T. Muramatsu, *J. Biol. Chem.*, 265 (1990) 10765.

23 M.D. Kuo, Y. Oda, J.S. Huang and S.S. Huang, *J. Biol. Chem.*, 265 (1990) 18749.

24 F.A. Marston, *Biochem. J.*, 240 (1986) 1.

25 R.J. Simpson and E.C. Nice, in A.R. Kerlavage (Editor), *The Use of HPLC in Receptor Biochemistry*, Alan R. Liss, New York, 1989, p. 204.

26 B. Grego, E.C. Nice and R.J. Simpson, *J. Chromatogr.*, 352 (1986) 359.

27 R.C. Beavis, T. Chaudhary and B.T. Chait, *Org. Mass Spec.*, 27 (1992) 156.

28 B.T. Chait, R. Wang, R.C. Beavis and S.B.H. Kent, *Proc. 40th ASMS Conference on Mass Spectometry and Allied Topics, Washington, DC, May 31–June 5, 1992*, p. 1939.

29 J.S. Cottrell, K.K. Mock and M.P. Stevens, *Proc. 39th ASMS Conference on Mass Spectometry and Allied Topics, Nashville, TN, May 1991*, p. 354.

30 J.J. Gorman, G.L. Corino and S.J. Mitchell, *Eur. J. Biochem.*, 168 (1987) 169.

31 M.J. O'Hare and E.C. Nice, *J. Chromatogr.*, 171 (1979) 209.

32 P.H. Corran, in R.W. Oliver (Editor), *HPLC of Macromolecules*, IRL Press, Oxford, p. 127.

33 J. Garnier, D.J. Osguthorpe and B. Robson, *J. Mol. Biol.*, 120 (1978) 97.

34 H. Marquardt, M.W. Hunkapiller, L.E. Hood and G.J. Todaro, *Science*, 223 (1984) 1079.

35 R.C. Czech, K.B. Clairmont, K.A. Yagaloff and S. Corvera, in M.B. Sporn and A.B. Roberts (Editors), *Peptide Growth Factors and Their Receptors I*, Springer, New York, p. 37.