Journal of Chromatography, 268 (1983) 197–206
Elsevier Science Publishers B.V., Amsterdam — Printed in The Netherlands

CHROM. 15,985

AMINO ACID SEQUENCE DETERMINATION BY GAS CHROMATOGRAPHY–MASS SPECTROMETRY OF PERMETHYLATED PEPTIDES

THE APPLICATION OF CAPILLARY COLUMNS

KEITH ROSE*, AMOS BAIROCH and ROBIN E. OFFORD
Département de Biochimie Médicale, C.M.U., 1 Rue Michel-Servet/9 Av. de Champel, 1211 Geneva 4 (Switzerland)

(Received March 29th, 1983)

SUMMARY

The application of capillary columns to the determination of the amino acid sequence of proteins and polypeptides by gas chromatography–mass spectrometry (GC–MS) of partial hydrolysates is described and discussed for the case of the $N^{\alpha,\varepsilon}$-trifluoroacetyl-N,O-permethyl derivatives. Retention indices are determined with the aid of a computer program. The mass spectra of methionine enkephalin obtained by GC–MS on a packed and on a capillary column are presented and a retention index calculated. Amounts of derivative corresponding to sub-nanomole amounts of peptides are sufficient to provide full amino acid sequence information. To assist in the assignment of retention indices above 4400, a mixture of *n*-alkanes from a commercial source was characterized by GC–MS up to *n*-$C_{54}H_{110}$, and was found to contain material up to an expected $C_{72}H_{146}$.

INTRODUCTION

The application of gas chromatography–mass spectrometry (GC–MS) to amino acid sequence determination involves partial hydrolysis of the polypeptide, which may be a fragment of a protein or even an intact small protein, to yield an overlapping set of smaller peptides. These peptides are rendered volatile by chemical treatment, separated by GC and their structures determined by MS. GC–MS analysis is able to provide sequence data rapidly and reliably from diverse regions of the original polypeptide. We have shown how GC–MS is especially useful when studying molecules that have undergone post-translation modification[1], using an idea that has been exploited more recently by others[2,3].

The application of capillary columns to the GC–MS analysis of peptides as their TMS-polyamino alcohol derivatives has been studied[4]. These derivatives have inferior mass spectral characteristics, but are more volatile than the permethylated derivatives[5]. Indeed, in order to exploit the mass spectral properties of permethylated peptides whilst remaining relatively free of volatility constraints, experiments have

0021-9673/83/$03.00 © 1983 Elsevier Science Publishers B.V.

been performed using high-performance liquid chromatography (HPLC)–MS[6]. The HPLC resolution was poor, and the least volatile substance studied was leucine enkephalin[6], which should in principle be amenable to GC. We believe that HPLC should be used for the separation of proteins and polypeptide fragments, rather than for complex partial hydrolysates that have undergone a derivatization aimed at reducing the polarity of their components to a minimum.

We have previously applied support-coated open-tubular (SCOT) capillary columns to the GC–MS analysis of partial hydrolysates of peptides as their $N^{\alpha,\varepsilon}$-trifluoroacetyl-N,O-permethyl derivatives[7]. Although useful results were obtained, packed columns remained the method of choice. Peptides of (approximately) four residues or more suffered unacceptable losses either on injection on to SCOT columns, or on transfer from the column to the mass spectrometer. SCOT columns also suffer from the disadvantage that they cannot be operated at very high temperature.

These problems were solved, theoretically at least, by the introduction of on-column injection and of wall-coated open-tubular (WCOT) capillary columns made of fused silica, which can be operated at high temperatures (350°C) and also inserted into the ion source of the mass spectrometer.

In this paper we describe the application of fused-silica WCOT capillary GC columns to the analysis of permethylated peptides. We show that the low-bleed, low-adsorption characteristics of such columns permit the successful analysis of compounds considered difficult[6], such as methionine enkephalin. In addition, the higher resolution (over that obtained with short packed columns) leads to more precise determinations of the Kováts retention indices, a measure which can be more useful than the amino acid composition[8].

EXPERIMENTAL

The GC–MS equipment consisted of a Carlo Erba Model 4161 gas chromatograph coupled to a Kratos MS 50 S mass spectrometer (Kratos, Urmston, Manchester, U.K.) via a direct coupling[9]. The mass spectrometer was operated in the low-resolution mode, with accelerating voltage 8 kV, electron beam energy 70 eV and source temperature 200°C. Data were acquired and processed by a DS 55 data system (Kratos). Gas chromatography was performed on various lengths of a variety of fused-silica columns (CPSil5, 0.12–0.4 μm coating, 0.31–0.34 mm I.D.; Chrompack, The Netherlands) unless otherwise stated. Helium was used as the carrier gas.

Samples of peptide derivative were applied in solution in chloroform (0.2–2 μl) using the on-column injector. Particular experimental conditions are given in the figure legends. GC experiments with flame-ionization detection were performed in a similar fashion on a Carlo Erba 4160 gas chromatograph or, for packed column work, on a Pye 104 instrument.

Derivatives were prepared as described previously[1].

RESULTS AND DISCUSSION

Choice of parameters

In most instances, the product of the action on a polypeptide of our preferred endopeptidase, subtilisin, is a mixture of peptides of 2–6 residues in length. This

mixture of compounds may thus contain some components close to or beyond the limits of volatility of a GC–MS analysis under reasonable conditions (column length, flow-rate, etc.). It is worthwhile to devote some effort to changing the operating conditions so as to be able to elute as many components as possible, so long as an acceptable degree of GC resolution is preserved. Each peptide (in particular the longer ones) that can be brought into the working range of the system as a result of such changes may contain valuable sequence information. Examples of possible changes would be shortening the column, increasing the flow-rate or, while taking care that sufficient sample capacity is preserved, the use of a thinner film of stationary phase. However, taken to extremes, such changes would be counterproductive, as the resolution would not be adequate and with wider GC peaks the sensitivity would be lower. What is required is a satisfactory compromise between the various parameters in Table I. The sensitivity of the mass spectrometer is also important if one is to avoid the necessity of overloading columns when scanning rapidly over a wide mass range; difficulty in obtaining complete spectra has been reported in the analysis of alkanes above n-$C_{36}H_{74}$ on a quadrupole instrument[10].

After many experiments in which the various parameters were systematically altered, we are able to recommend a useful working compromise (Table I). We shall discuss each parameter in turn.

The choice of a column I.D. of 0.31–0.34 mm permits the use of the unmodified Carlo Erba on-column injector, and provides a higher sample capacity than a narrow-bore (0.25 mm) column. When a direct coupling is used, 0.31–0.34 mm I.D. columns lead to lower MS source pressures than would be produced with wide-bore (0.5 mm) columns.

We chose CPSil 5 (100% methylsilicone) as it is thermally stable (particularly on fused silica) and of low polarity. SE-54, a slightly more polar phase (95% methylsilicone, 4% phenylsilicone, 1% vinylsilicone) shows noticeably higher retention of permethylated peptides (see Table II). Whilst this effect tends to resolve more components of a complex mixture by spreading them over a wider range of retention indices, it also tends to shift the less volatile peptides beyond the range accessible under reasonable GC conditions. The 100% methylsilicone phase is therefore to be preferred.

A film thickness of 0.12–0.14 μm, taken with a column I.D. of 0.31–0.34 mm, provides a sample capacity that corresponds to about 200 pmol of a tetrapeptide. This is more than is required for flame-ionization detection when determining retention indices (see below), and is adequate for GC–MS work. However, if full spectra are to be obtained from peptides that are produced in low yield by the enzyme used for the partial hydrolysis, it may be necessary to exceed the loading of 200 pmol for the major components. A column with a film thickness of 0.4 μm was found to have a sample capacity of about 1 nmol and is therefore to be preferred. A film thickness of 1.2 μm, whilst providing a large sample capacity, is much too thick as only very small peptides may be eluted under reasonable conditions of column length and flow-rate.

After trials with column lengths between 1.2 and 26 m, we found 5–8 m to be the most useful. Such lengths provide a resolution much greater than that obtainable on short packed columns (see below) yet permit peptides of low volatility (retention indices greater than 4600) to be eluted. With helium as the carrier gas, inlet pressures of 0.15–0.3 kg/cm^2 lead to optimum GC resolution for flame-ionization operation

TABLE I

PARAMETERS FOR CAPILLARY COLUMN GC OF SAMPLES OF LOW VOLATILITY

Parameter	*Value**	*Advantages***	*Disadvantages***	*Range examined*	*Preferred compromise*
Column I.D.	Narrow	Higher resolution than wider columns; lower MS source pressure (direct coupling)	Lower sample capacity; non-standard injection if <0.3 mm I.D.	0.31–0.34 mm	0.31–0.34 mm
Polarity of stationary phase	Low	Lower elution temperature	Less selective	CPSil 5, SE-54	CPSil 5
Thickness of stationary phase film	Thin	Lower elution temperature	Lower sample capacity; shorter column life; risk of adsorption	0.12, 0.14, 0.4, 1.2 μm	0.4 μm
Column length	Short	Lower elution temperature; economical	Inferior resolution	1.2–26 m***	5 m
Inlet pressure/carrier gas flow-rate	High	Lower elution temperature; shorter analysis time	Higher MS source pressures (if direct coupling used)	0.1–2.5 kg/cm^2 [§] 0.7–5 ml/min [§§]	0.15–0.3 kg/cm^2 [§] 0.7–2 ml/min [§§]

* Qualitative value of the parameter favouring the analysis of samples of low volatility.

** Advantages and disadvantages of each qualitative value for low volatility analyses by GC–MS.

*** 5 m was the shortest column tried for GC–MS work; shorter columns were tried with the flame-ionization detector.

[§] Inlet pressure (above ambient), atmospheric outlet (flame-ionization detector) operation.

[§§] Vacuum outlet (GC–MS) operation.

TABLE II

RETENTION INDICES OF CERTAIN PEPTIDES AS THEIR $N^{\alpha,\varepsilon}$-TRIFLUOROACETYL-N,O-PERMETHYL DERIVATIVES DETERMINED ON CAPILLARY COLUMNS COATED WITH TWO DIFFERENT STATIONARY PHASES

The error (computational and due to measurement) is $< 1\%$. Amino acid sequences are given in the one-letter code.

Peptide	*Retention index on CPSil 5*	*Retention index on SE 54*
TS	1772	1813
LM	2005	2054
NT	2131	2200
VQ	2152	2214
DF	2235	2294
WL	2560	2630
GTF	2582	2665
KY	2813	2888
MNT	2959	3033
SKY	3380	3461
WLM	3420	3519
VQW	3632	3753
VQWL	4282	4363

with the column outlet at atmospheric pressure, and columns of 5–8 m × 0.31–0.34 mm I.D. We recommend a length of 5 m for a film thickness of 0.4 μm and 8 m for a film thickness of 0.12 μm.

In order to obtain optimum GC resolution under GC–MS conditions (in which the outlet is under vacuum), it is necessary to reduce the gas flow-rate by operating at subambient inlet pressures. This may be done simply, with no instrumental modifications; the inlet pressure gauge reads zero under these conditions, but adjustments to the gas regulator on the gas chromatograph are seen to affect the vacuum reading on the mass spectrometer ion gauge. Flow-rates, expressed in ml/min, are measured at ambient pressure at the pump outlets. For example, the resistance to flow offered by an 8 m × 0.34 mm I.D. column leads to a flow-rate of about 5 ml/min (measured at ambient pressure at the pump outlet) when its inlet pressure is just above ambient (a pressure excess of 0.12 kg/cm^2). If the flow-rate is reduced to 0.7 ml/min, closer to the optimum under vacuum outlet conditions, the inlet pressure falls below ambient. The combined effect of increased GC resolution and improved MS sensitivity due to lower source pressures then leads to an increase in sensitivity of a factor of four. The advantages of sub-ambient inlet operation have been described in detail[11]. To avoid pronounced solvent tailing, we prefer to inject at an inlet pressure just above ambient and to regulate the flow-rate during the first few minutes of operation. This fact, together with the finite cycle time of the mass spectrometer (normally 2.5–3.5 sec) leads to some uncertainty in the measurement of retention indices, and we prefer to use the flame-ionization detector trace for all precise measurements. Although this implies dividing the sample into two portions, the flame-ionization detector measurements only require amounts of derivative that correspond to 100 pmol or less of the polypeptide; at our present stage of development of the derivatization technique, this is not a significant fraction of the total amount needed.

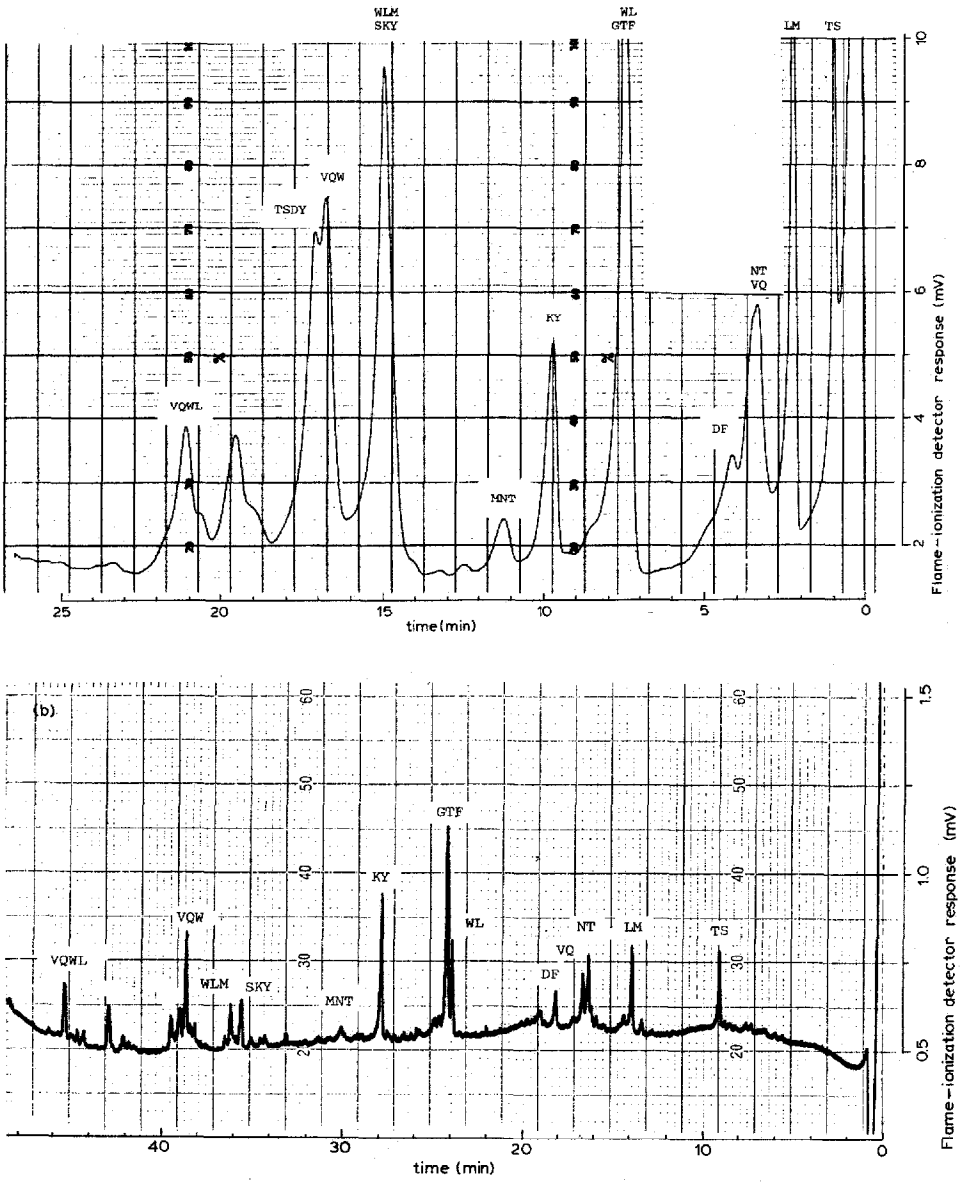

Fig. 1. (a) Packed column analysis. Flame-ionization detector trace of a subtilisin digest of glucagon after $N^{\alpha,\varepsilon}$-trifluoroacetylation and N,O-permethylation. No derivatization of arginine prior to permethylation, essential if arginine-containing peptides are to yield volatile derivatives, was performed. Amino acid sequences, in the one-letter code, denote the elution position of the corresponding peptides. Chromatographic conditions: 30 cm × 2 mm I.D. column filled with 1% OV-1 on Chromosorb W HP (100–120 mesh) (Chrompack); nitrogen carrier gas, flow-rate 15 ml/min; oven temperature, 150°C, programmed at 6.5°C/min to 310°C; 11-nmol sample injected in 1 μl of chloroform. (b) Capillary column analysis. Flame-ionization detector trace of a portion of the same derivatized subtilisin digest of glucagon, here analysed on a short capillary column. Chromatographic conditions: 5 m × 0.34 mm I.D. fused-silica WCOT column coated with a 0.12 μm film of CPSil 5 (Chrompack); helium carrier gas, inlet pressure 0.25 kg/cm^2; oven temperature, 80°C, 1 min isothermal, programmed at 4°C/min to 320°C; *ca.* 300 pmol of sample injected in 0.3 μl of chloroform.

Comparison of analyses on packed and capillary columns

Fig. 1 compares the flame-ionization detector traces given by the same derivatized partial hydrolysate when run on a packed column and on a short capillary column. Even though the capillary column performance was compromised in this instance by operation at a high flow-rate, the increased resolution is still apparent. This increased resolution, which facilitates mass spectral interpretation in a GC–MS experiment as peaks are more likely to exist as single, separated components (compare Fig. 1a and b), permits a more precise determination of the retention index for the same reason. For these retention index measurements, a mixture of *n*-alkanes (those with an even number of carbon atoms between C_{18} and C_{44}, with C_{38} missing) is injected under a chosen set of conditions. The sample is afterwards co-injected with a small amount of alkane mixture under the same conditions. The alkanes are easily picked out by their regularity after superposition of the traces. After amino acid sequence analysis of the component peptides in a GC–MS experiment under similar but not identical conditions (vacuum outlet), retention indices are calculated for identified components from the flame-ionization detector trace using an interactive

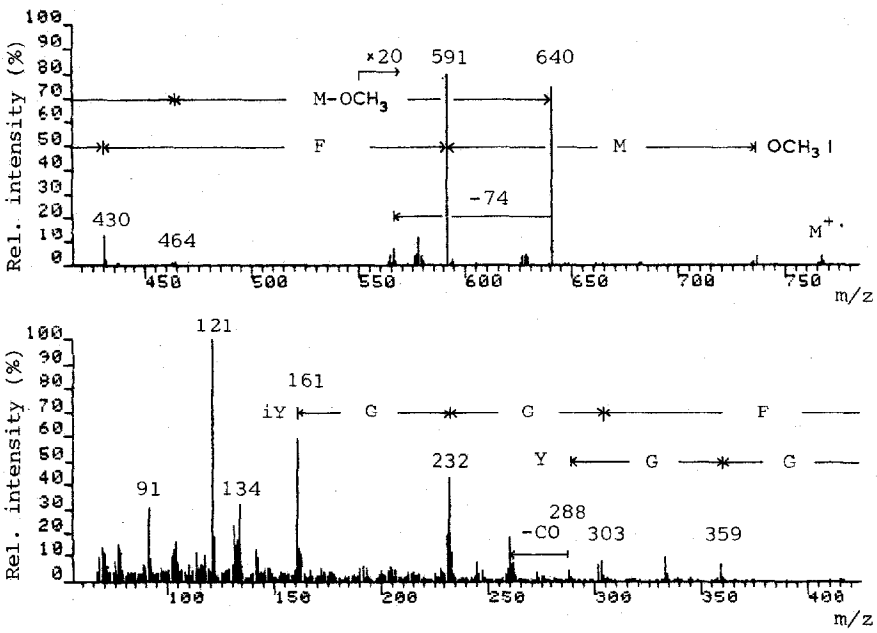

Fig. 2. Mass spectrum of methione enkephalin, as its N^{α}-trifluoroacetyl-N,O-permethyl derivative, obtained by GC–MS using a capillary column. Chromatographic conditions: 5-m fused-silica column (see Fig. 1b); helium carrier gas; oven temperature, 80°C, 1 min isothermal, programmed at 4°C/min to 310°C. An amount of sample corresponding to *ca.* 6 nmol of enkephalin was injected in 1 μl of chloroform. Two series of sequence ions are present, both indicating the entire sequence, Tyr–Gly–Gly–Phe–Met. A small degree of under-methylation is responsible for the signals at *m/z* 577 and 626. The first series, denoted by the one-letter code, commences at *m/z* 288. The second series, corresponding to in-chain cleavage at the Tyr residue and denoted iY..., commences at *m/z* 161. The ion at *m/z* 134 corresponds to $\overset{+}{H}NCH_3{=}CHCH_2C_6H_5$ and indicates that Phe is present but not in the N-terminal position[13]. Ions at *m/z* 91 and 121 are due to the side-chain of Phe and Tyr, respectively.

computer program that fits the alkane data to a curve using the cubic-spline method[12]. The interactive program is written in Fortran IV (Data General revision IX) and copies are available from the authors on request. The computational error was estimated by omitting each alkane in turn from the calibration calculation and calculating its retention index according to its position on the recorder trace. The maximum error found in the index was 0.54%. Error due to inaccurate measurement of the recorder trace (±0.5 mm) is about 0.3% at retention index 1800 at a chart speed of 0.5 cm/min, diminishing to about 0.1% at retention index 4400 at a chart speed of 1 cm/min. The global error (computational and measurement) is thus expected to be less than 1%.

We were unable to obtain supplies of pure *n*-alkanes above *n*-$C_{44}H_{90}$. Fortunately, low-molecular-weight polythene is available to extend the data points beyond retention index 4400. A sample of "M.Wt. 700 polythene" (Polymer Laboratories, Church Stretton, Shropshire, U.K.) was analysed by GC coupled both with a flame-ionization detector and with a mass spectrometer. The mixture was found, by co-elution with standards up to *n*-$C_{44}H_{90}$ and by GC–MS analysis up to *n*-$C_{54}H_{110}$[9], to consist, as expected, of a homologous series of *n*-alkanes beginning at about *n*-$C_{30}H_{62}$ (smaller amounts of earlier members are present) and continuing up to *n*-$C_{54}H_{110}$. Indeed, the flame-ionization detector trace showed that the series continued beyond *n*-$C_{54}H_{110}$ up to at least *n*-$C_{72}H_{146}$. Although the later members of the series were not characterized, it is reasonable, because of the regularity of the series and the

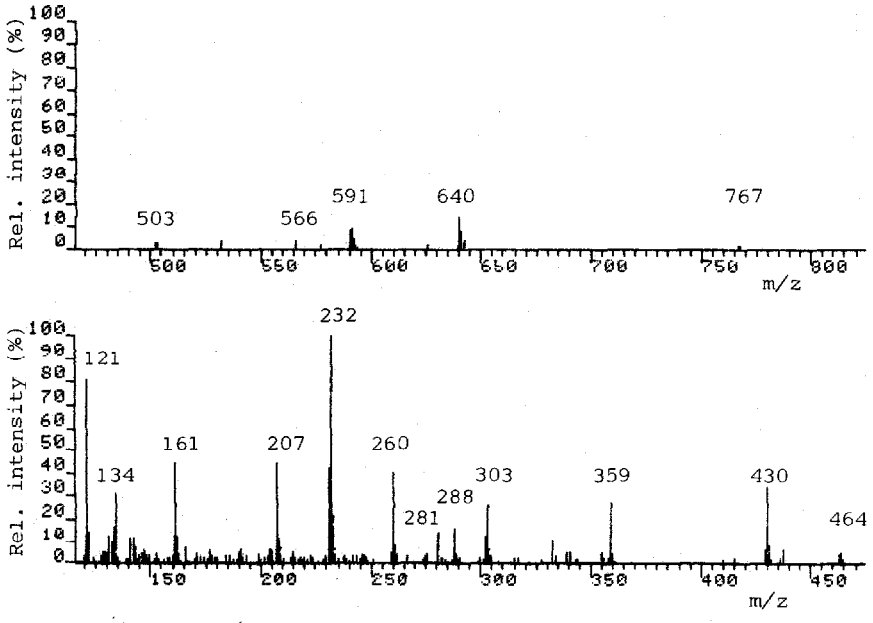

Fig. 3. Mass spectrum of methionine enkephalin, as its N$^{\alpha}$-trifluoroacetyl-N,O-permethyl derivative, obtained by GC–MS using a packed column. Chromatographic conditions as in Fig. 1a, except that helium was used as the carrier gas and the upper temperature was 340°C. A different GC–MS system, which has been previously described[1], was used for this experiment. The ion series *m/z* 207, 281, 341, 355, 429, 503 is due to column bleeding. An amount of derivative corresponding to approximately 50 nmol of peptide was injected in 1 µl of chloroform.

origin of the material, to assume that they correspond to the higher *n*-alkanes. By mixing this material with analytical standards of n-$C_{18}H_{38}$–n-$C_{44}H_{90}$ alkanes, it is possible to establish a retention index scale from 1800 to 7200. Indeed, by adding "M.Wt. 1000 polythene" from same supplier, the scale may be extended to 8200 or higher.

The mass spectrum of methionine enkephalin obtained by GC–MS on a capillary column is presented in Fig. 2, to be compared with the mass spectrum of the same material obtained by GC–MS on a packed column (Fig. 3). The sequence is easily deduced in either case, but the capillary analysis is to be preferred as the background due to column bleed is less prominent. The peptide derivative eluted from a CPSil5 capillary column with a retention index of 4653.

Amounts of material

Because capillary columns produce little bleeding, adsorb little of the sample irreversibly and give narrow peaks, it is possible to use them to analyse sub-nanomole amounts of material by GC–MS. Amounts of derivative corresponding to sub-nanomole amounts of peptide yield full mass spectra. For example, 200 pmol of the derivative of Val–Gly–Ser–Glu was sufficient (Fig. 4). Such results have higherto been obtained by derivatizing very much larger amounts (at least 100 nmol) and then injecting the appropriate fraction. We find that a minor modification of the permethylation procedure allows us to carry out the derivatization successfully down to the 2 nmol level even in the presence of substantial amounts of sodium dodecyl sulphate[14].

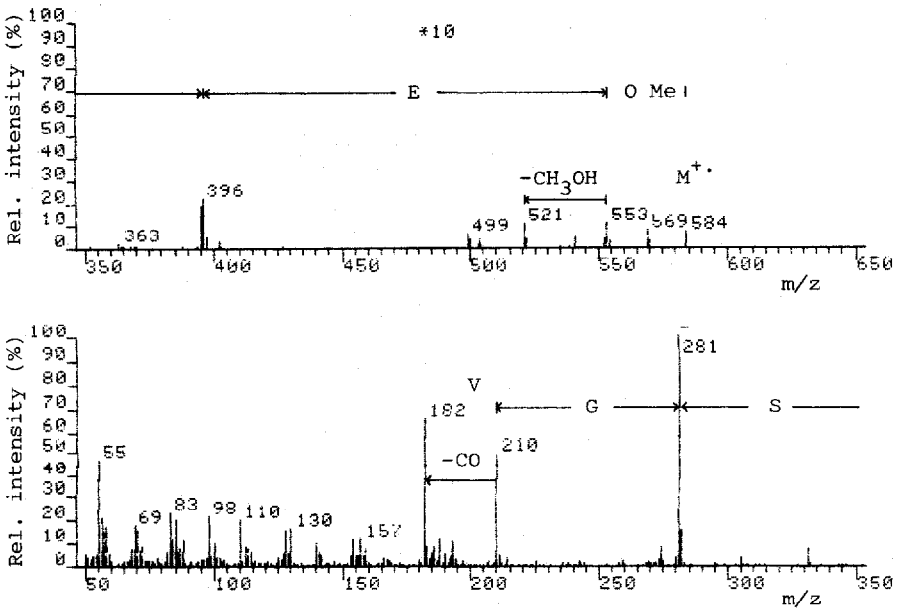

Fig. 4. Mass spectrum of the tetrapeptide Val–Gly–Ser–Glu, as its N^{α}-trifluoroacetyl-N,O-permethyl derivative, obtained by GC–MS using a fused-silica WCOT column (8 m × 0.34 mm I.D.) with a 0.12-μm film of CPSil 5. The oven temperature was 120°C, 1 min isothermal, programmed at 4°C/min to 310°C. Apart from during the injection, the flow-rate of helium carrier gas was 0.7 ml/min. An amount of derivative corresponding to 200 pmol of peptide was injected in 0.2 μl of chloroform.

CONCLUSION

Certain capillary columns can be operated under conditions that permit the successful analysis by GC–MS of peptides as their permethylated derivatives. Chromatographic resolution is improved, which facilitates mass spectral interpretation, retention indices are more precise and mass spectra have lower background signals than when packed columns are used. If a sensitive mass spectrometer and an efficient derivatization procedure are used, capillary column GC–MS offers a rapid approach to sequence analysis at the low-nanomole level.

ACKNOWLEDGEMENTS

We thank the Fonds National Suisse de la Recherche Scientifique for a grant to purchase the equipment. We also acknowledge the financial assistance of Biogen SA.

REFERENCES

1 K. Rose, J. D. Priddle, R. E. Offord and M. P. Esnouf, *Biochem. J.*, 187 (1980) 239–243.
2 S. A. Carr and K. Biemann, *Biomed. Mass Spectrom.*, 7 (1980) 173–178.
3 S. A. Carr, P. V. Hauschka and K. Biemann, *J. Biol. Chem.*, 256 (1981) 9944–9950.
4 H. J. Stan and M. Banz, *Anal. Chem. Symp. Ser.*, 4 (1980) 53–61.
5 H. Nau, *J. Chromatogr.*, 121 (1976) 376–382.
6 T. J. Yu, H. Schwartz, R. W. Giese, B. L. Karger and P. Vouros, *J. Chromatogr.*, 218 (1981) 519–533.
7 J. D. Priddle, K. Rose and R. E. Offord, in A. Frigerio (Editor), *Advances in Mass Spectrometry in Biochemistry and Medicine*, Vol. II, Spectrum Publications, New York, 1976, pp. 477–485.
8 P. J. W. Schuyl, B. van de Graaf and E. Izeboud, *Advan. Mass Spectrum.*, 8B (1980) 1549–1553.
9 K. Rose, *J. Chromatogr.*, 259 (1983) 445–452.
10 G. Dielman, S. Meier and U. Rapp, *J. High Resolut. Chromatogr. Chromatogr. Commun.*, 2 (1979) 343–350.
11 C. A. Cramers, G. J. Scherpenzeel and P. A. Leclercq, *J. Chromatogr.*, 203 (1981) 207–216.
12 J. H. Ahlberg, E. N. Nilson and J. L. Walsh, *The Theory of Splines and Their Applications*, Academic Press, London, 1967.
13 K. Rose, *D. Phil. Thesis*, Oxford, 1979.
14 K. Rose, M. G. Simona and R. E. Offord, *Biochem. J.*, in press.