Journal of Chromatography, 301 (1984) 425–440
Elsevier Science Publishers B.V., Amsterdam — Printed in The Netherlands

CHROM. 16,926

SEQUENCE ANALYSIS OF DERIVATIZED PEPTIDES BY HIGH-PERFORMANCE LIQUID CHROMATOGRAPHY–MASS SPECTROMETRY

T. J. YU, H. A. SCHWARTZ, S. A. COHEN, PAUL VOUROS* and B. L. KARGER*
Barnett Institute of Chemical Analysis and Materials Science, and Department of Chemistry, Northeastern University, Boston, MA 02115 (U.S.A.)
(Received May 25th, 1984)

SUMMARY

N-Acetyl-N,O,S-permethylated derivatives of oligopeptides were analyzed by high-performance liquid chromatography–mass spectrometry (HPLC–MS) using a moving belt interface. A heated-gas nebulizer was employed for sample deposition, thus permitting the effective use of a water–methanol gradient covering the range from 5% to 95% water at mobile phase flow-rates of 0.5 ml/min. We demonstrate in this paper that it is possible to sequence the octapeptide derived from the C-chain of glucagon by HPLC–MS analysis of a permethylated enzymatic hydrolysate of this peptide using overlap information from the mass spectral patterns. Moreover, it is shown that peptides not readily amenable to analysis by gas chromatography–MS can be analyzed using this approach. Preliminary results suggest that N-acetyl-methyl ester derivatives of oligopeptides may in specific cases also be a useful alternative for HPLC–MS analysis of complex oligopeptide mixtures.

INTRODUCTION

In a previous paper, the approach of peptide analysis by on-line liquid chromatography–mass spectrometry (LC–MS) using N-acetyl-N,O,S-permethylated peptides was illustrated[1]. This strategy followed logically the successful off-line chromatography–MS[2–4], as well as gas chromatography (GC)–MS[5] approaches of others using the same derivative. With a heated-gas inlet for nebulization and rapid solvent removal, it was possible to employ gradient elution in reversed-phase liquid chromatography (RPLC) with a moving-belt interface. Promising initial results were reported in the previous paper. Thus, chemical ionization (CI), using isobutane or ammonia, yielded sequence determining fragmentation at both the N- and C-terminals. Examples of good on-line mass spectra included leucine enkephalin which produced an $[M + 1]^+$ ion of m/z 696.

As a continuation of that work, this paper presents other examples of N-acetylated-N,O,S-permethylated peptides, including analysis of an enzymatic digest of an octapeptide. The possibility of using the N-acetylated methyl ester derivative of a peptide for LC–MS sequence analysis is also explored. The results of this paper

0021-9673/84/$03.00

lend further evidence of the potential of the LC–MS approach for sequencing of oligopeptides.

MATERIALS AND METHODS

Materials

Glucagon carboxyl terminal octapeptide was purchased from Chemical Dynamics (S. Plainfield, NJ, U.S.A.), and protease Type VII (subtilisin BPN') and chymotrypsin (TLCK treated) were from Sigma (St. Louis, MO, U.S.A.). Hydrogen chloride, technical grade, was obtained from Matheson (Gloucester, MA, U.S.A.). The sources of all other materials were as described in our previous paper[1].

Derivatization and enzymatic hydrolysis

The acetylation and permethylation procedures described by Leclerq and Desiderio[2] were employed. Peptides containing arginine were treated with hydrazine to convert arginine to ornithine before acetylation, as described by Thomas *et al.*[6]. Esterification was performed by methanolic hydrogen chloride (3 *N*). The samples were refrigerated when not in use.

Proteolytic digestions were carried out in 0.1 *M* aqueous ammonium bicarbonate. The digests consisted of 1:50 (w/w) ratio of enzyme to peptide (approximately 2 mg total peptide) in a total volume of 0.5 ml. The reactions were allowed to proceed for 4 h at room temperature, frozen, and freeze dried twice prior to derivatization for LC–MS. The progress of the digestion was monitored at various times by analyzing 2-μl aliquots by RPLC using a linear gradient (0–40% organic) in which the aqueous solvent was 10 m*M* phosphoric acid and acetonitrile the organic solvent.

Instrumentation

The mass-spectrometric approaches, using a moving belt interface, have been described previously, including all operational conditions[1]. All mass spectra were recorded in the CI mode using isobutane as reagent gas. Careful adjustment of the heated-gas nebulizer for solvent evaporation[7] permitted successful LC–MS operation at a flow-rate of 0.5 ml/min from acetonitrile–water (5:95, v/v); hence, essentially over the whole solvent range in RPLC. The high-performance liquid chromatographic (HPLC) equipment was also as previously listed[1]. Chromatography was performed on a Supelcosil LC-8 column, 15 cm × 4.6 mm I.D. (Supelco, State College, PA, U.S.A.).

RESULTS AND DISCUSSION

N-Acetyl permethylated derivatives

This portion of the study represented a continuation of previously reported work[1] to evaluate the usefulness of permethylated derivatives of peptides for analysis by HPLC–MS. More specifically, the goal was to assess whether the sequence of a polypeptide could be determined from the overlap data provided by the permethylated oligopeptides analyzed during HPLC–MS of enzymatic hydrolysates. This approach is essentially similar to the GC–MS procedure which relies on the use of volatile O-trimethylsilyl polyaminoalcohol derivatives of peptides[7] or of N-trifluoroacetyl-N,O,S-permethylated peptides[5].

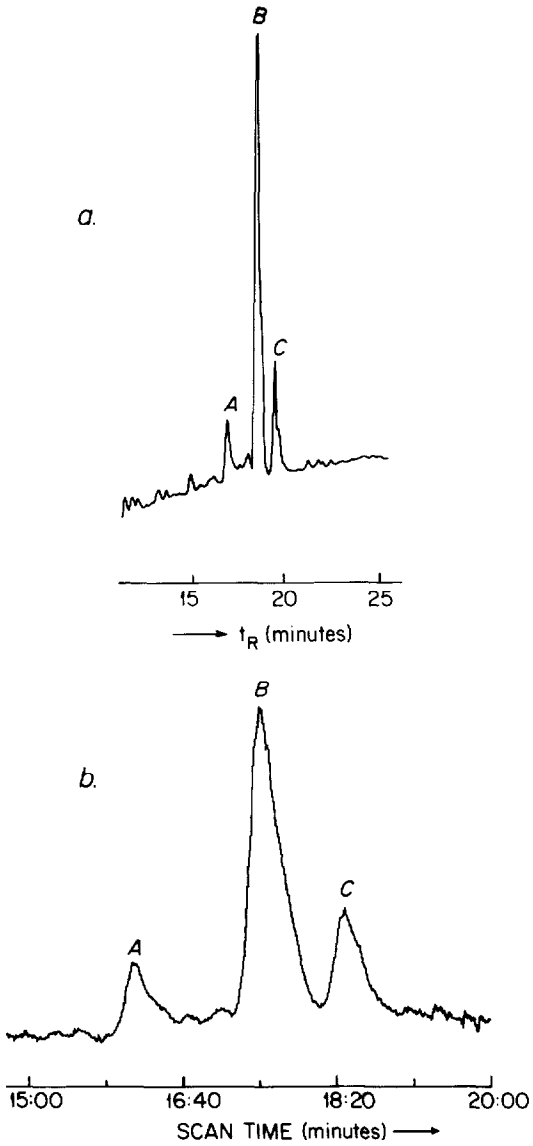


Fig. 1. HPLC–UV (a) and HPLC–MS (b) chromatograms of permethylated Met-enk. Peaks: A = Ac-Me · Tyr-Me · Gly-Me · Gly-Me · Phe-OMe; B = Met-enk [Ac-Me · Tyr-Me · Gly-Me · Gly-Me · Phe-Me · Met-OMe]; C = C-methylate Met-enk.

As indicated earlier, the obvious advantage of HPLC–MS lies in the significantly lower volatility requirements for HPLC as opposed to GC. An example of the potential advantage is illustrated in Figs. 1 and 2 which show the results obtained with methionine-enkephalin (Met-enk), a pentapeptide (Tyr-Gly-Gly-Phe-Met) whose O-TMS-polyaminoalcohol derivative has a calculated GC retention index of 4320[8], *i.e.*, a range not routinely accessible by GC. A chromatogram (HPLC–UV) of Met-enk following permethylation is shown in Fig. 1a and compared to the

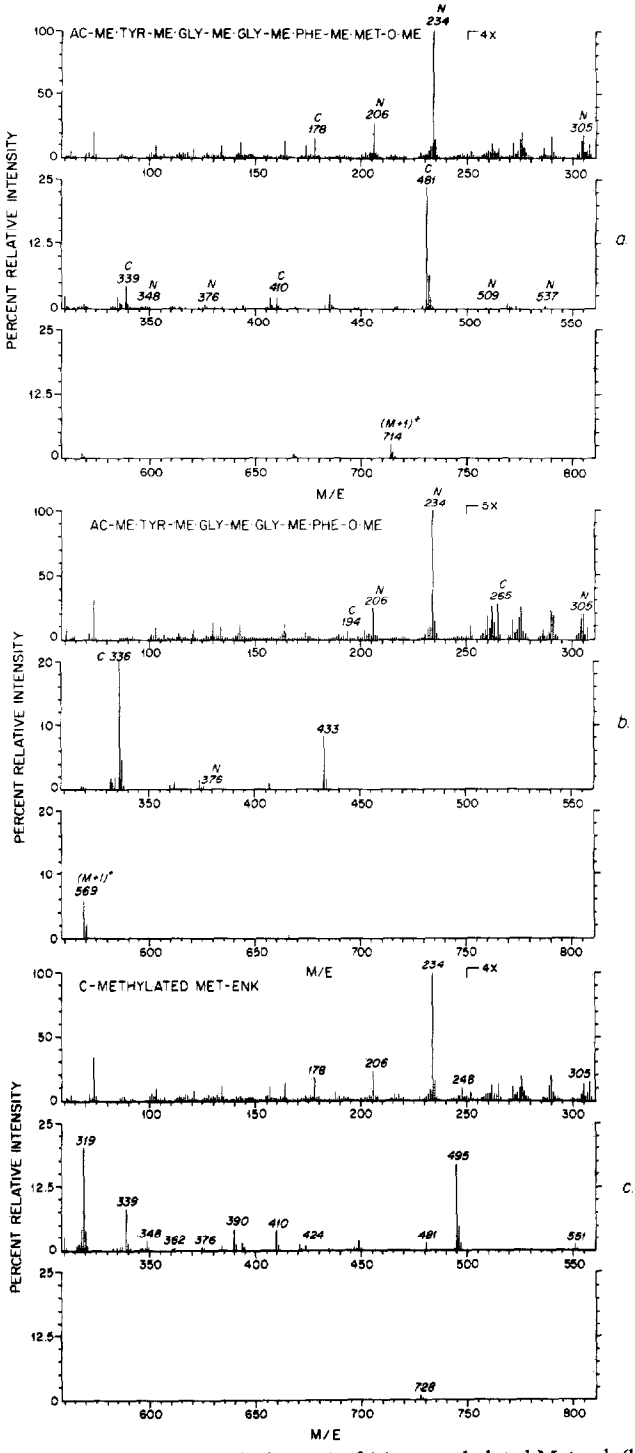
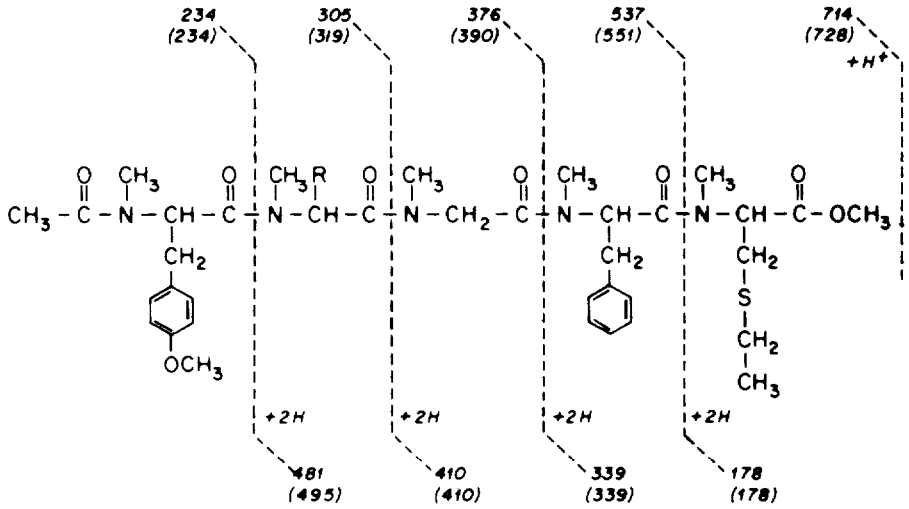


Fig. 2. Mass spectra (CI-isobutane) of (a) permethylated Met-enk (b) Ac-Me · Tyr-Me · Gly-Me · Gly-Me · Ph-OMe and (c) C-methylated Met-enk from HPLC–MS chromatogram of Fig. 1b.

HPLC–MS total-ion chromatogram (Fig. 1b). Chromatographic resolution during HPLC–MS is to a large extent maintained as a result of the incorporation of a heated-gas nebulizer for solvent deposition. Indeed, improved nebulizer designs over that used to generate the data shown in Fig. 1 have shown little or no increase in band broadening from the interface in combination with either normal-bore[9] or microbore[10] HPLC columns.

The major peak (B) in Fig. 1 was identified from its mass spectrum as N-acetyl-permethylated Met-enk (Fig. 2a). The two smaller peaks were identified as the N-acetyl-permethylated tetrapeptide Tyr-Gly-Gly-Phe (peak A, Fig. 2b), presumably produced by hydrolysis of Met-enk during the derivatization process and C-methylated Met-enk (peak C, Fig. 2c). Moreover, the isobutane-CI mass spectra of the three major peaks of Fig. 1 exhibit fragment ions reflecting the amino acid sequence in the respective oligopeptides. For example, the spectrum of Met-enk (Fig. 2a) shows a molecular adduct ion at *m/z* 714 and the indicated sequence determining N- and C-terminal ions. Adduct ions associated with the attachment of a *tert.*-butyl group (57 a.m.u.) to a fragment ion are also occasionally observed as exemplified by the ion at *m/z* 433 [376 + 57] in the spectrum of the tetrapeptide (Fig. 2b). Furthermore, comparison of the spectrum of Met-enk (Fig. 2a) with its C-methylated counterpart (Fig. 2c) allows determination of the position of C-methylation[11]. For example, examination of the N-terminal sequence ions suggests that C-methylation involves mainly the glycine residue attached to the N-terminal tyrosine (note the 14 a.m.u. shift of the ion at *m/z* 305, Scheme 1). This is further corroborated from inspection of the C-terminal ions, particularly the shift of the intense m/z 481 peak in Fig. 2a to *m/z* 495 in Fig. 2c.



Scheme 1. Fragmentation (CI-isobutane) of N-acetyl-permethylated Met-enk and C-methylated derivative of the Gly residue attached N-terminal Tyr.

The results of Met-enk indicate that oligopeptide mixtures can be analyzed by permethylation and HPLC–MS. In order to assess the general applicability of this approach, we next examined the enzymatic hydrolysates of the octapeptide Phe-Val-

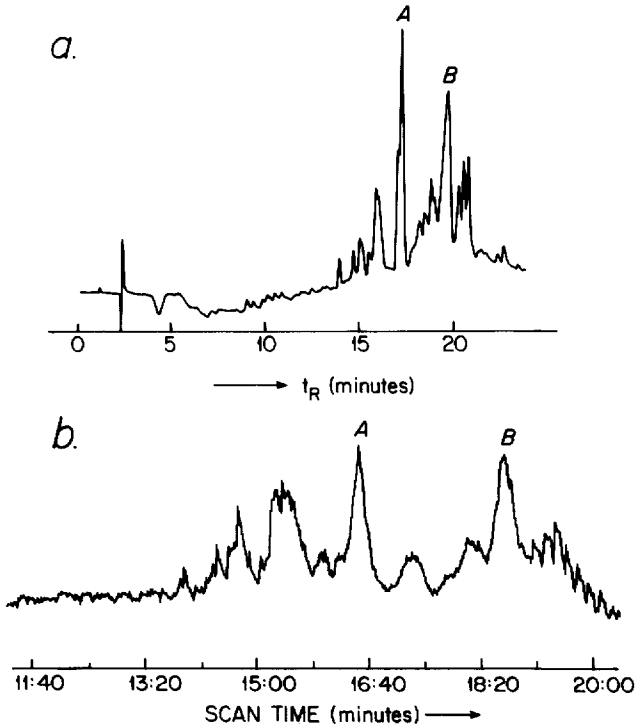
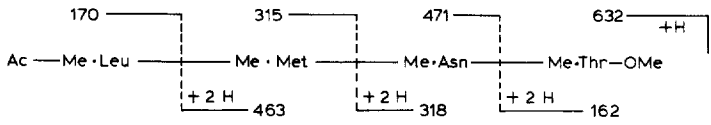


Fig. 3. HPLC–UV (a) and HPLC–MS (b) chromatograms of permethylated peptides from the chymotrypsin hydrolysate of the C-chain of glucagon.

Glu-Trp-Leu-Met-Asn-Thr, derived from the C-terminus of glucagon, using two different enzymes. Figs. 3a and b show the HPLC–UV and HPLC–MS chromatograms, respectively, of the hydrolysates using chymotrypsic digest of the peptide followed by permethylation. Two peptides resulting from a cleavage between tryptophan and leucine were the expected products. Accordingly, peaks A and B in Fig. 3 were identified from their mass spectra as those of the permethylated derivatives of Leu-Met-Asn-Thr (Fig. 4a) and Phe-Val-Glu-Trp (Fig. 4b). The peak at m/z 632 in Fig. 4a corresponds to the mass of the protonated molecular ion of the tetrapeptide Ac-Me · Leu-Me · Met-Me · Asn-Me · Thr-O · Me. Furthermore, N- and C-terminal fragment ions, as indicated on the spectrum, provide the structural information consistent with the sequence on Scheme 2.



Scheme 2. Fragmentation (CI-isobutane) of N-acetyl-permethylated Leu-Met-Asn-Thr (peak A in Fig. 3a and b).

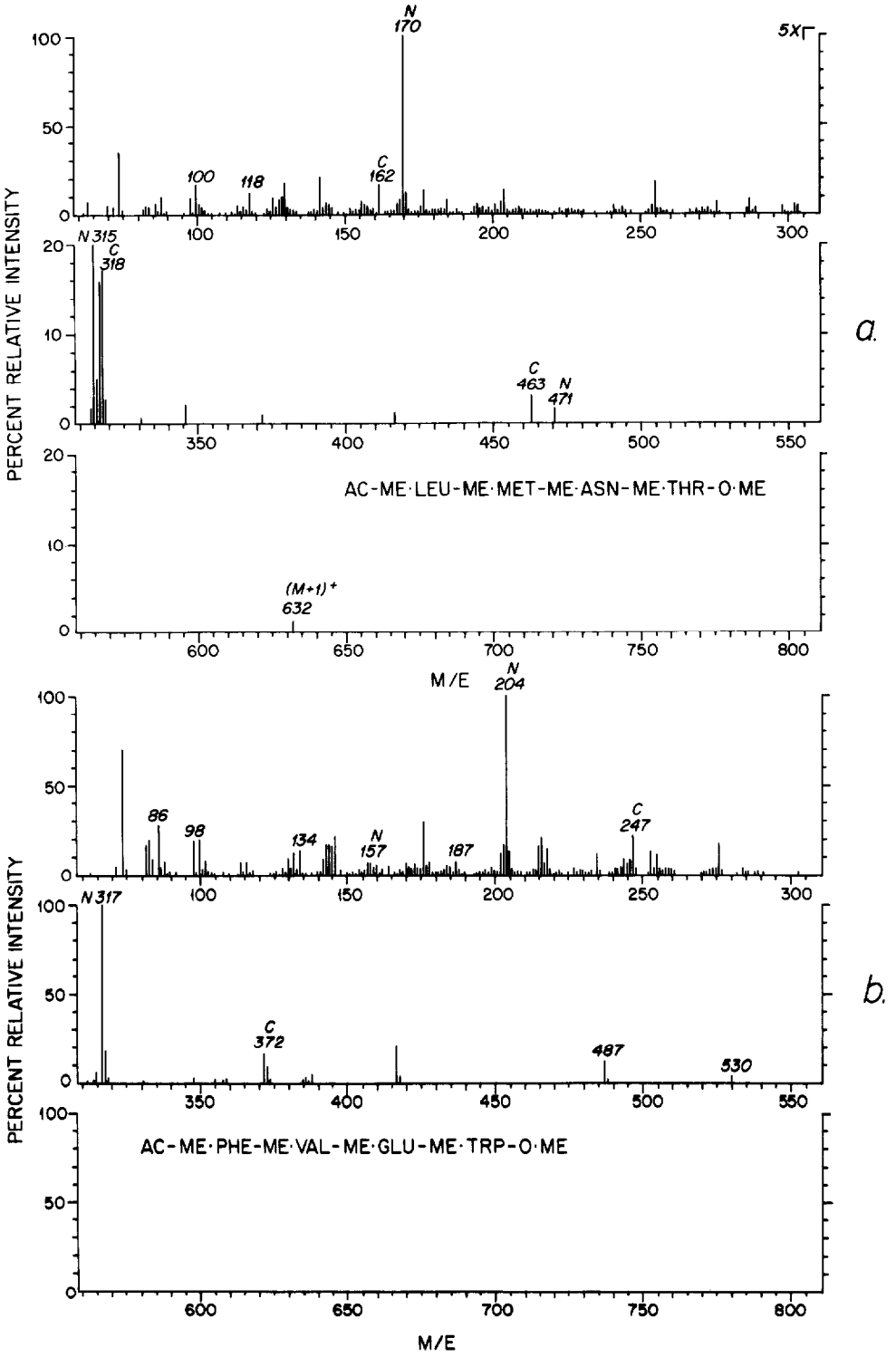


Fig. 4. Mass spectra (CI-isobutane) of (a) Ac-Me · Leu-Me · Met-Me · Asn-Me · Thr-OMe, peak A in Fig. 3 and, (b) Ac-Me · Phe-Me · Val-Me · Glu-Me · Trp-OMe, peak B in Fig. 3.

Somewhat less well-defined is the mass spectrum of peak B in the chromatogram of the glucagon hydrolysate (Fig. 4b). However, despite the absence of a molecular adduct ion peak of *m/z* 721, identification as the tetrapeptide Phe-Val-Glu-Trp is still possible by combining the information from N- and C-terminal ions. For example, the peaks at *m/z* 204 and 317 are indicative of a N-terminal permethylated Phe-Val sequence while C-terminal ions at *m/z* 247 and *m/z* 372 correspond to tryptophan and the Glu-Trp minus CH_3OH sequence, respectively (Scheme 3). Moreover, the identity of the amino acid residues in the peptide represented by peak B in the chromatograms of Fig. 3 is confirmed from the masses of imine ions at *m/z* 86 (Val, a), *m/z* 98 [(Glu-CH_3OH, b)], *m/z* 134 (Phe, c) and *m/z* 187 (Trp, d).

It is apparent from the above discussion that permethylation of the chymotrypsic hydrolysate and analysis of the mixture by HPLC–MS identified the two tetrapeptides comprising the original octapeptide. However, analysis of a second hydrolysate of the glucagon C-chain produced from a subtilisin enzyme digest provided for a more definitive determination of the octapeptide sequence. The HPLC–UV and HPLC–MS chromatograms of the latter digest following N-acetyl-permeth-

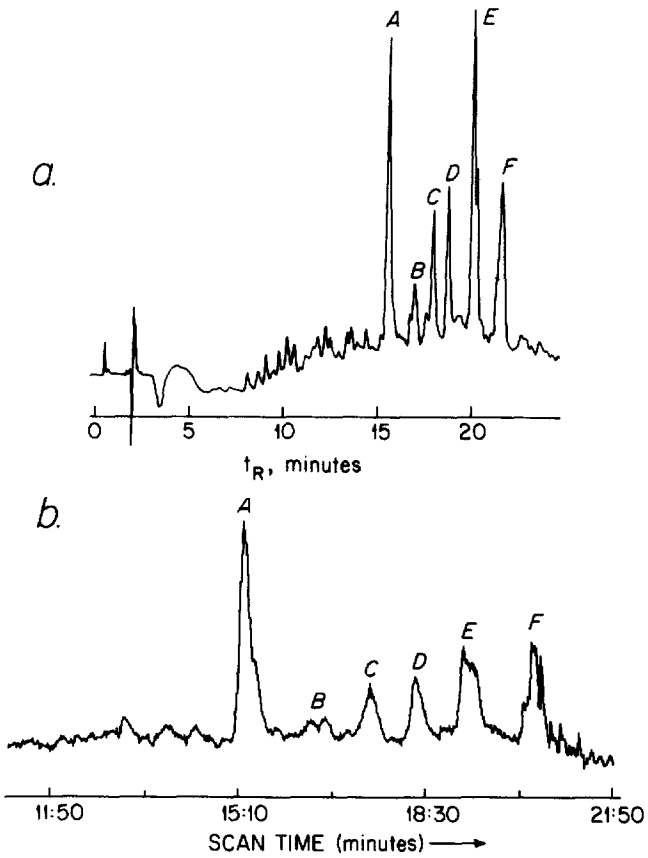
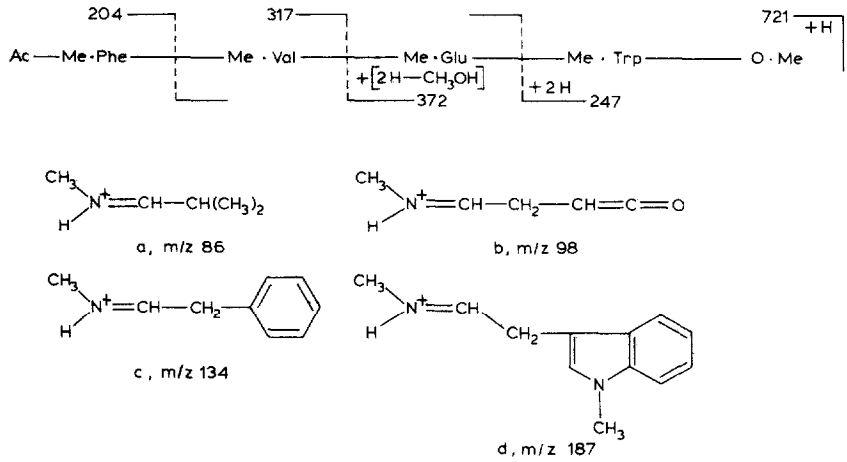


Fig. 5. HPLC–UV (a) and HPLC–MS (b) chromatograms of permethylated peptides from the subtilisin hydrolysate of the C-chain of glucagon. Peak F = Ac-Me · Trp-Me · Leu-Me · Met-OMe.



Scheme 3. Fragmentation (CI-isobutane) of N-acetyl-permethylated Phe-Val-Glu-Trp.

ylation are shown in Figs. 5a and b, respectively. Again, chromatographic fidelity for the HPLC–MS total-ion chromatogram may be noted from the comparison.

A pertinent chromatographic component is peak F which was identified from its mass spectrum as the N-acetyl-permethylated derivative of Trp-Leu-Met (Fig. 6). While no $[M + H]^+$ ion was observed, the fragment ions from the N- and C-ter-

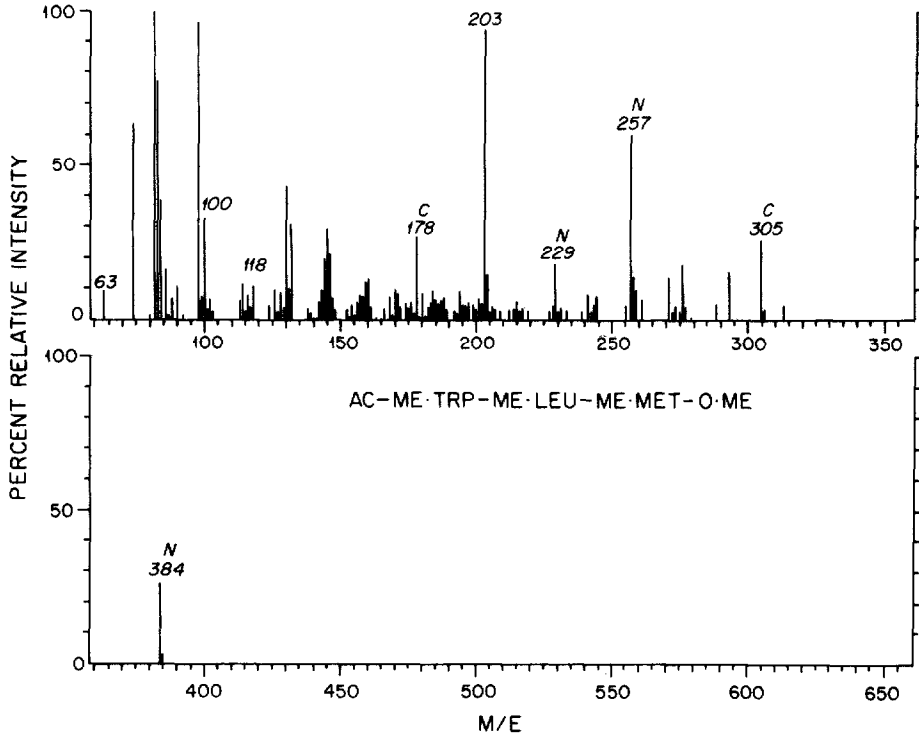
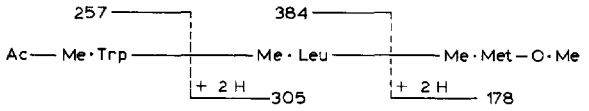


Fig. 6. Mass spectrum (CI-isobutane) of Ac-Me · Trp-Me · Leu-Me · Met-OMe, peak F in Fig. 5.

minals allowed definitive identification of the sequence of the tripeptides shown in Scheme 4. Moreover, from the overlap of Trp-Leu-Met it was possible to reconstruct the structure of the original octapeptide. Thus, these results demonstrate the potential of HPLC–MS, using a moving belt interface, for analysis of peptide mixtures following N-acetyl permethylation and for the determination of the structures of polypeptides using "overlap" information from hydrolysates.



Scheme 4. Fragmentation (CI-isobutane) of N-acetyl-permethylated Trp-Leu-Met.

N-Acetyl-methyl ester derivatives

While the results described above indicate the potential of standard permethylation procedures for the analysis of peptides by HPLC–MS, some drawbacks are apparent. Of primary concern is the relative complexity of the derivatization reaction which often results in the formation of multiple side products due, in part, to C-methylation or to lower yields when precautions are taken to minimize the occurrence of such side reactions. We thus examined the utility of HPLC–MS of N-acetyl-methyl

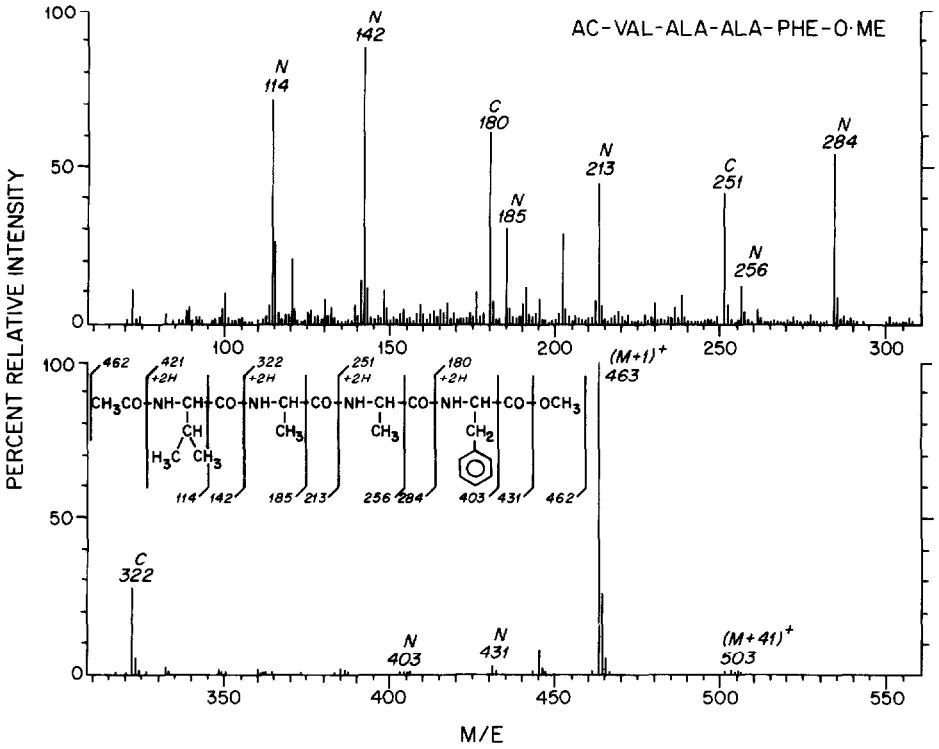


Fig. 7. Mass spectrum (CI-isobutane) and fragmentation pattern of N-acetyl-methyl ester derivative of Val-Ala-Ala-Phe.

ester derivatives of oligopeptides, with the anticipation that this simpler two-step derivatization procedure would produce a cleaner product in a higher yield than the more elaborate permethylation process.

N-perfluoroacyl-O-methyl ester derivatives of dipeptides have been used by Seift *et al.*[12] for the analysis of peptide mixtures by GC–MS. The best structural information was often obtained by combining the data from electron impact and CI mass spectra. Arpino and McLafferty[13] employed N-acetyl-O-methyl ester derivatives for the analysis of oligopeptides by HPLC–MS using a direct-liquid introduction interface. Due to the limited amount of data obtained in this latter study, it was not possible to ascertain the value of the chemical ionization mass spectra for the determination of amino acid sequence in oligopeptides. Therefore, in conjunction with our evalution of the N-acetylation-methyl esterification procedure for HPLC–MS, we examined the CI mass spectra of several standard peptides containing up to five amino acid residues in order to determine the value of these derivatives for sequence determination by CI–MS. Some representative examples are discussed below.

The CI-isobutane mass spectrum of the N-acetyl-methyl ester derivative of the tetrapeptide Val-Ala-Ala-Phe is shown in Fig. 7 along with a summary of its fragmentation pattern. Structurally characteristic N- and C-terminal fragment ions are observed throughout, along with a well-defined $[M + H]^+$ molecular adduct ion. A similarity with the CI fragmentation pattern of permethylated derivatives as origi-

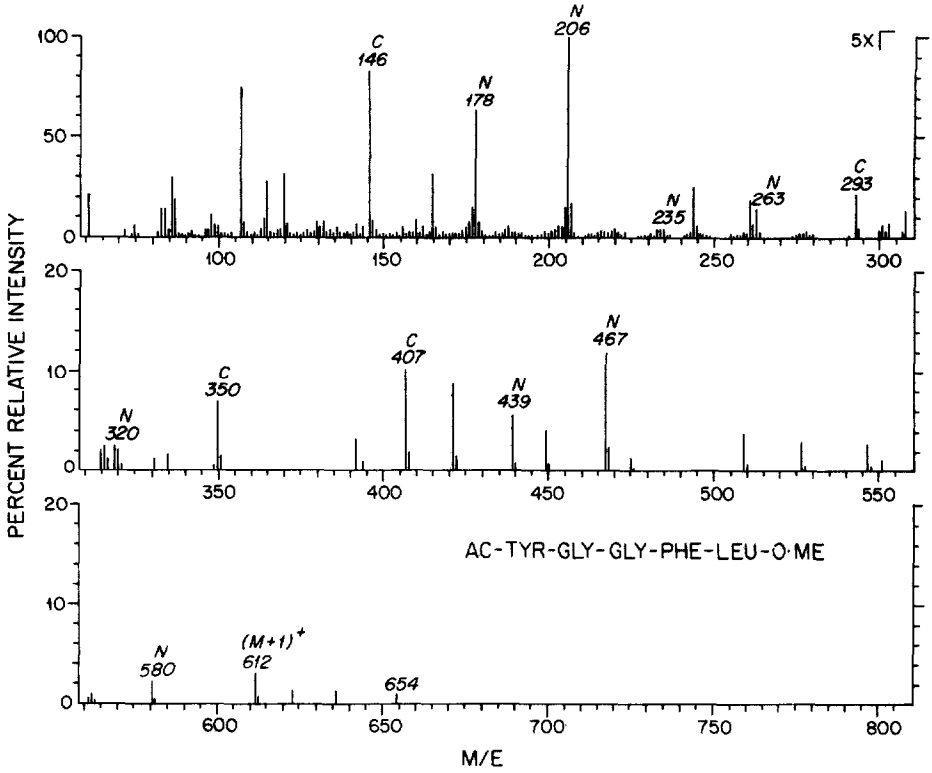
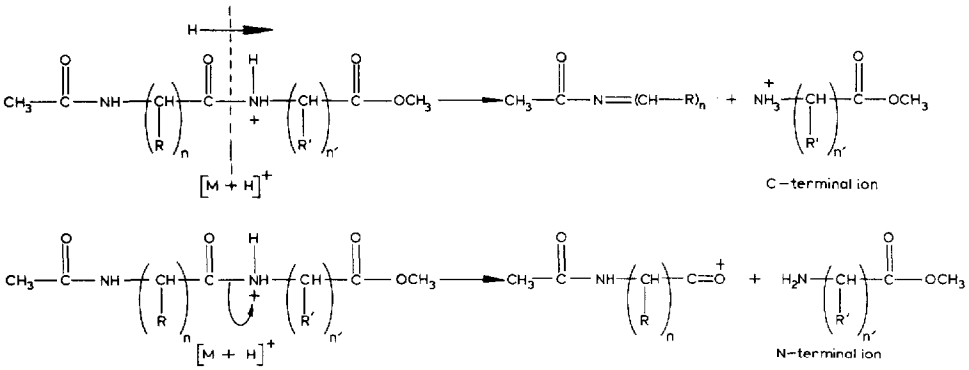


Fig. 8. Mass spectrum (CI-isobutane) of N-acetyl-methyl ester derivative of Tyr-Gly-Gly-Phe-Leu.

nally proposed by Field and co-workers[4,14] may also be noted. Namely, fragmentation leading to a C-terminal ion involves cleavage of the peptide bond of the protonated molecular ion of the peptide and is accompanied by a hydrogen transfer from the remaining neutral moiety to form a quaternary ammonium ion. On the other hand, formation of the N-terminal ion involves a homolytic cleavage of the peptide bond and charge localization on the carbonyl oxygen. These processes are summarized in Scheme 5.



Scheme 5. Fragmentation mechanisms leading to N- and C-terminal ions in the CI-isobutane mass spectra of N-acetyl-O-methyl ester derivatives of peptides.

In general, similar results were obtained with the other peptides shown in Figs. 8 and 9. The spectrum of the N-acetyl-methyl ester derivative of the pentapeptide Tyr-Gly-Gly-Phe-Leu (Fig. 8) again displays all the significant N- and C-terminal ions along with the $[M + H]^+$ peak. This consistency was retained with the histidine-containing tripeptide, Gly-His-Gly (Fig. 9a) and the dipeptide Trp-Gly (Fig. 9b).

In order to determine the usefulness of N-acetyl-methyl ester derivatives of peptides for HPLC–MS a synthetic mixture of eight peptides was examined. The HPLC–UV and HPLC–MS chromatograms of the derivatized mixture are shown in Figs. 10 and 11 respectively. Comparison with the corresponding chromatograms obtained previously with the permethylated derivatives of the same mixture[1], shows a clearer definition of the chromatographic peaks in the N-acetyl-methyl ester derivatives. This is primarily due to the absence of side reaction products which often result in multiple peaks for each peptide (*e.g.*, C-methylation in the permethylated derivatives). Therefore, in that respect, it may be inferred that N-acetyl-methyl ester derivatives may provide a useful alternative to permethylation for HPLC–MS. Two peptides in the mixture (Val-Tyr-Val and Phe-Pro) were found to co-elute and are unresolved from their HPLC–UV and HPLC–MS total-ion chromatograms (Figs. 10 and 11), but are readily deconvoluted from the mass chromatograms (Figs. 12c and h).

In conclusion, these data demonstrate the potential of HPLC–MS for sequencing of oligopeptides. It has been shown that use of a heated-gas nebulizer for deposition of the HPLC effluent onto the moving belt interface allows effective use of gradient elution for analysis of derivatized oligopeptide mixtures preserving the chromatographic fidelity of the system. These results are further reinforced by im-

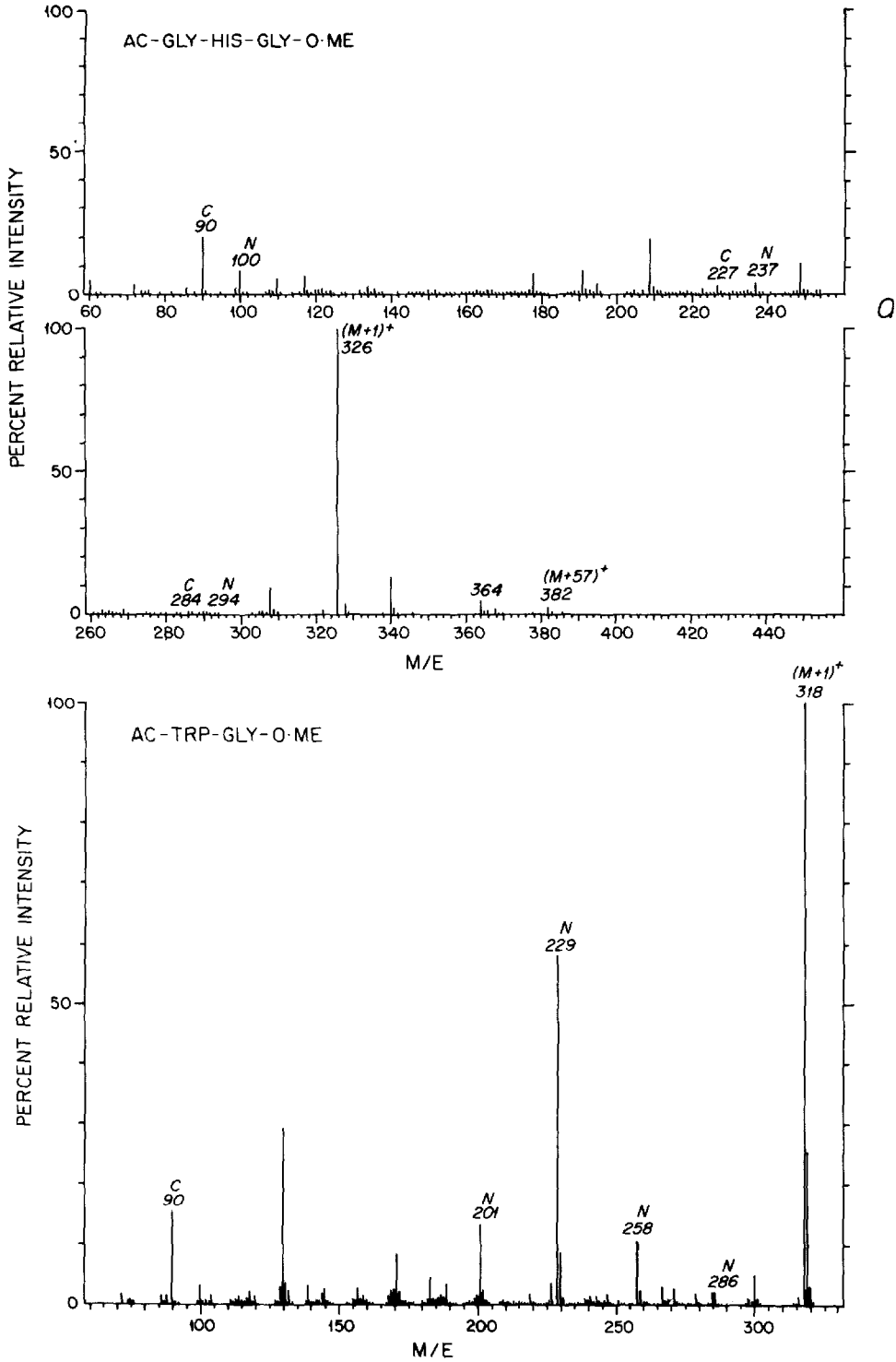


Fig. 9. Mass spectra (CI-isobutane) of N-acetyl-methyl ester derivatives of (a) Gly-His-Gly and (b) Trp-Gly.

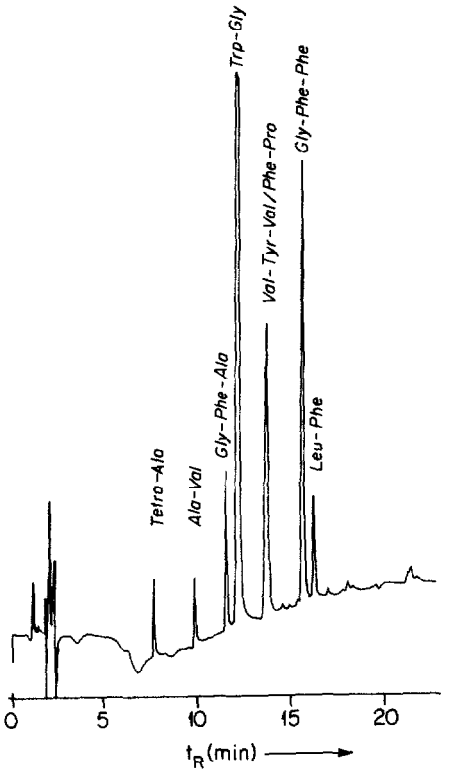


Fig. 10. HPLC–UV chromatogram of N-acetyl-methyl ester derivatives of a mixture of eight synthetic peptides.

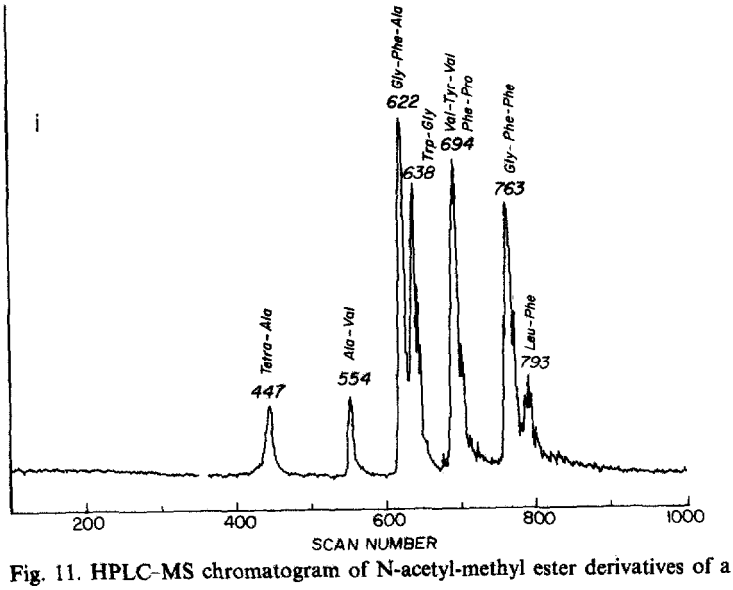


Fig. 11. HPLC–MS chromatogram of N-acetyl-methyl ester derivatives of a mixture of eight synthetic peptides.

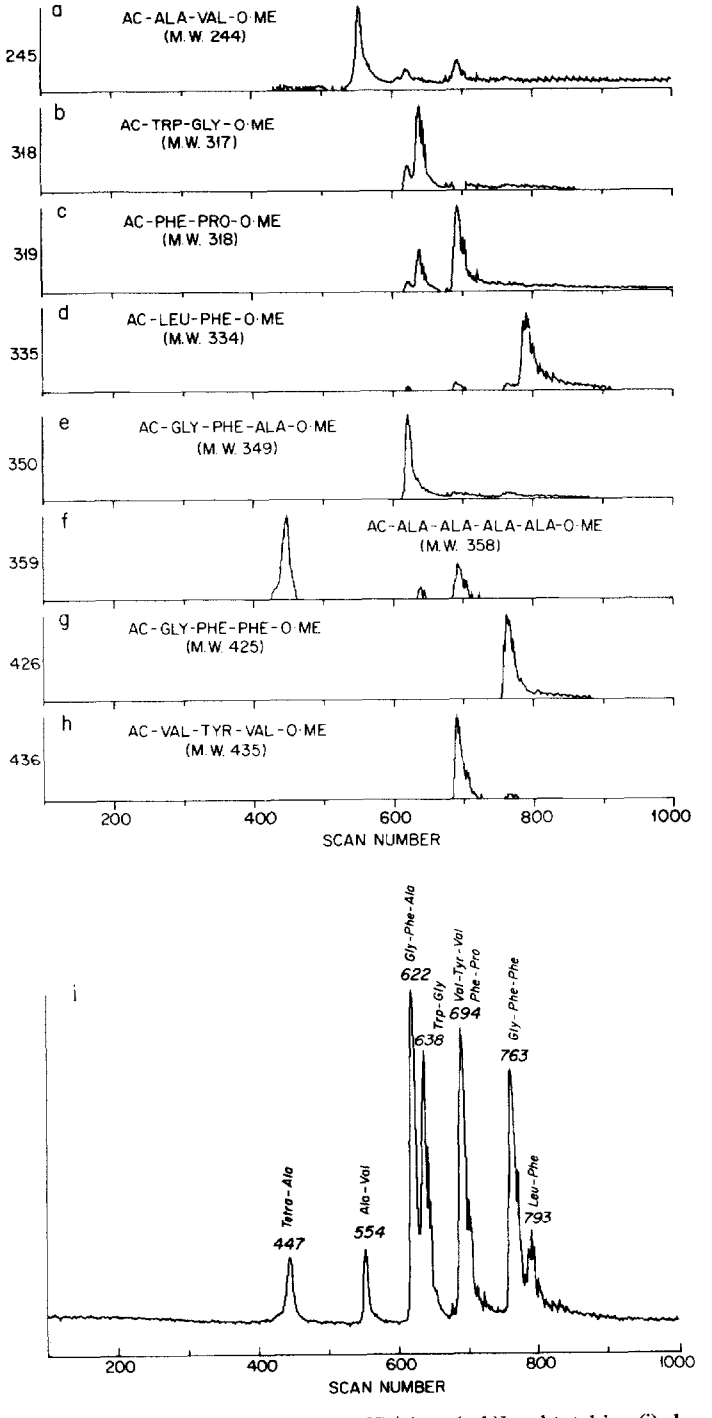


Fig. 12. Mass chromatograms [$(M + H)^+$ ions (a–h)] and total-ion (i) chromatograms of N-acetyl-methy ester derivatives of synthetic peptide mixture of Fig. 10.

proved nebulizer designs[9,10]. Larger peptides can be sequenced by analysis of the permethylated derivatives of oligopeptides produced from their enzymatic hydrolysis. On the other hand, it has been shown that, at least from the examples investigated, it may also be possible to conduct HPLC–MS analysis of oligopeptide mixtures based on the formation of N-acetyl-methyl esters.

The purpose of this work has been to demonstrate the principles of HPLC–MS for sequencing using precolumn derivatization. Accordingly, no attention has been paid to sample size requirements and relatively large amounts of peptide have been sequenced. It is to be noted that low-ng sensitivity for full spectrum scans has been shown in LC–MS using the moving belt interface[10] and indeed in favorable cases pg sensitivity[9]. With proper attention to precolumn reaction and use of improved nebulizer designs, ng sensitivity should be possible for this approach to sequencing. It may also be noted that peptide standards have been used in this work. Studies with samples from synthetic mixtures or extracts are necessary to validate the method. It is anticipated that this procedure will complement other approaches involving MS[15], Edman sequencing[16] and DNA sequencing[17].

ACKNOWLEDGEMENTS

This work was supported by a grant from the National Institute of Health, GM 15847 (BLK). Contribution No. 195 from the Barnett Institute of Chemical Analysis and Materials Science.

REFERENCES

1 T. J. Yu, H. Schwartz, R. W. Giese, B. L. Karger and P. Vouros, *J. Chromatogr.*, 218 (1981) 519.
2 P. A. Leclerq and D. M. Desidero, *Anal. Lett.*, 4 (1971) 305.
3 H. R. Morris, *Nature (London)*, 286 (1980) 447.
4 M. Mudgett, D. V. Bowen, J. J. Kindt and F. H. Fields, *Biomed. Mass Spectrom.*, 2 (1975) 254.
5 K. Rose, J. D. Priddle, R. E. Offord and M. P. Esnouf, *Biochem. J.*, 187 (1980) 239.
6 D. W. Thomas, B. C. Das, S. D. Gero and E. Lederer, *Biochem. Biophys. Res. Commun.*, 32 (1968) 519.
7 K. Biemann, *Pure Appl. Chem.*, 50 (1978) 169.
8 H. Nau and K. Biemann, *Anal. Biochem.*, 73 (1976) 139.
9 M. J. Hayes, E. P. Lankmayr, P. Vouros, B. L. Karger and J. M. McGuire, *Anal. Chem.*, 55 (1983) 1745.
10 M. J. Hayes, H. E. Schwartz, P. Vouros, B. L. Karger, A. D. Thruston, Jr. and J. M. McGuire, *Anal. Chem.*, 56 (1984) 1229.
11 T. J. Yu, B. L. Karger and P. Vouros, *Biomed. Mass Spectrom.*, 10 (1983) 633.
12 W. F. Seifert, Jr., R. E. McKee, C. F. Beckner and R. M. Caprioli, *Anal. Biochem.*, 88 (1978) 169.
13 P. J. Arpino and R. W. McLafferty, in F. C. Nachod, J. J. Ackerman and E. W. Randall (Editors), *Determination of Organic Structures by Physical Methods, Vol. 6*, Academic Press, New York, 1976, pp. 1–89.
14 M. Mudgett, J. A. Sogan, D. V. Bowen and F. H. Field, *Advan. Mass Spectrom.*, 7 (1978) 1056.
15 R. G. Cooks, K. L. Busch and G. L. Glish, *Science*, 222 (1983) 273.
16 M. W. Hunkapiller, R. M. Hewick, W. J. Dreyer and L. E. Hood, in M. Elzinga (Editor), *Methods in Protein Analysis*, Humana Press, Clifton, NJ, 1982, p. 77.
17 L. E. Post and M. Nomura, *J. Biol. Chem.*, 255 (1980) 4660.