Journal of Chromatography, 197 (1980) 31–41
Elsevier Scientific Publishing Company, Amsterdam — Printed in The Netherlands

CHROM. 12,906

ANALYSIS OF DIPEPTIDE MIXTURES BY THE COMBINATION OF ION-PAIR REVERSED-PHASE HIGH-PERFORMANCE LIQUID CHROMATOGRAPHIC AND GAS CHROMATOGRAPHIC–MASS SPECTROMETRIC TECHNIQUES*

SHEN-NAN LIN*, LAURENCE A. SMITH and RICHARD M. CAPRIOLI
Analytical Chemistry Center and the Department of Biochemistry and Molecular Biology, The University of Texas Medical School, Houston, TX 77025 (U.S.A.)*

(First received December 27th, 1979; revised manuscript received April 23rd, 1980)

SUMMARY

Reversed-phase high-performance liquid chromatography using aliphatic carboxylic acids as surfactants was used to separate basic, acidic and neutral dipeptides. The enhancement of the capacity factors of basic dipeptides *versus* the chain length of carboxylic acid followed sigmoidal curves. The capacity factor of dipeptides fell precipitously with addition of salts until it decreased to a value which was unaffected by the addition of more salts. For the identification of dipeptides produced by the hydrolysis of proteins by proteases (dipeptidylaminopeptidases), gas chromatographic–mass spectrometric techniques were used for acidic and neutral dipeptides. Basic dipeptides could be positively identified either by normal chromatographic procedures or by collecting the suspected peaks for further characterization by alternate techniques.

INTRODUCTION

Unprotected peptides are difficult to separate by adsorption chromatography because of their polar nature. The separation of short-chain peptides has traditionally been accomplished by ion-exchange chromatography[1–3]. Only recently, reversed-phase high-performance liquid chromatography (HPLC) has been widely employed for the separation of amino acids and peptides. Retention is attributed to hydrophobic interaction between the solute and the hydrocarbonaceous functions covalently bound to the stationary surface. Peptides composed mainly of neutral amino acids are eluted in the order of increasing hydrophobicity[4]. Introduction of a polar substance such as alcohols in the eluent or a polar functional group in the solute reduces the capacity factor of a given compound. The ionization of certain groups in the solute molecule, which can be brought about by adjusting the pH of the eluent, has a similar effect. Chromatographic data reported for short-chain peptides are generally applicable only to neutral peptides[4,5]. Dipeptides composed of acidic or basic amino acids as well as those neutral dipeptides consisting of residues such as serine, glycine, proline, threo-

* This work is submitted by L.A.S. in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Graduate School of Biomedical Sciences, University of Texas Health Science Center at Houston.

0021-9673/80/0000–0000/$02.25

nine etc. are only poorly retarded in the hydrocarbonaceous columns. Therefore the same set of conditions usually employed for the analysis of neutral peptides[4,6] is inadequate to obtain the resolution required to analyze polar peptides. Ion-pair chromatography with the eluent containing ionic surfactant has been employed extensively to analyze ionizable materials[11,12], but only limited data has been reported for the analysis of polar peptides[4]. Since methods are sought for the determination of the primary structures of polypeptides through the analysis of enzyme released dipeptides from polypeptide substrates[13–15], HPLC techniques were employed in the analysis of basic dipeptides which present some difficulties in the gas-chromatographic–mass spectrometric (GC–MS) method normally used in the dipeptide analysis. The polypeptides ribonucelase S and val-4 angiotensin III were chosen as substrates to illustrate the methodology.

EXPERIMENTAL

Reagents and instrumentation

In all experiments glass-distilled organic solvents (Burdick & Jackson labs., Muskegon, MI, U.S.A.) and water were used. Dipeptides were purchased from either Sigma (St. Louis, MO, U.S.A.) or Vega (Tucson, AZ, U.S.A.). Waters HPLC instrumentation was used (Milford, MA, U.S.A.), including a U6K universal injector, a Model 6000 solvent delivery system, a Model 450 variable wavelength detector (monitored at 214 nm) and a 300 × 4 mm I.D. μBondapak C_{18} column. The gas chromatograph–mass spectrometer was a Finnigan Model 3200/6000 GC–MS data system, equipped with a solid probe sampling unit. The GC column was 1.5 m × 2 mm I.D. U-shaped glass column packed with 3% OV-1 on 80–100 mesh Gas-Chrom Q (Applied Science Labs. State College, PA, U.S.A.). Sample injection into the gas chromatograph occurred with the injector at 230°C and a linear temperature program from 100 to 300°C at 10°C/min. The glass jet separator was held at 260°C and the ion source was operated at 100°C. The electron impact source was operated at an ionization energy of 70 eV.

Enzymatic hydrolysis of polypeptides

About 5–10 nmole of ribonuclease S-peptide or val-4-angiotensin III were dissolved in a solution containing 30 μl 5% (v/v) lutidine, 79 μl water, 49 μl 0.5% (v/v) acetic acid and 47 μl 0.01 *N* HCl. The pH of the mixture was adjusted to 6.5 with acid or lutidine, if necessary, to ensure a peak enzyme activity[15]. After equilibration at 37°C for 15 min, digestion was initiated by adding 3–5 units of dipeptidylaminopeptidase (DAP) I/IV per μmole of substrate. The incubation period lasted for 2–3 h at 37°C. The mixture was frozen immediately and lyophilized. Finally, 40 μl of glass distilled water was added and 1–2 μl of the mixture was used for HPLC analysis.

Derivatization

The HPLC eluent containing the dipeptides was transferred to a Reacti-vial equipped with a PTFE-lined screw cap and lyophilized to dryness. After addition of approximately 200 μl of dry methanol, the mixture was cooled in a dry ice–ethanol bath, and then 40 μl of thionyl chloride was slowly added. The solution was heated in a constant temperature block at room temperature for 20 min. After removal of the

reagents in vacuo, 100 μl of perfluoropropionic anhydride (PFPA; Pierce, Rockford, IL, U.S.A.) was added and allowed to react at room temperature for 15 min. The PFPA was removed *in vacuo*, the residue was dissolved in 50 μl of dry acetone, and 5 μl of the solution was used for GC/MS analysis. Dipeptides containing arginyl residues required additional derivatization of the guanidine moiety for GC–MS analysis. Conversion of the arginine residue to the (4,6-dimethylpyrimid-2-yl)-ornithine derivative was accomplished using the method of Bacon *et al.*[16].

RESULTS AND DISCUSSION

HPLC separation of dipeptides

The separation of fourteen dipeptides, differing widely in their polarities, by reversed-phase HPLC with a mobile phase consisting of 0.038% (w/v) octanoic acid, 0.002% (w/v) pentanoic acid and 5% (v/v) butanol in water is shown in Fig. 1. Basic dipeptides Arg–Asp, Lys–Ser, His–Ala, Arg–Val, Lys–Lys and His–Phe, are well separated from acidic and neutral dipeptides. However, if a neat aqueous mobile phase is used instead, both the acidic and basic dipeptides are only poorly retarded and the neutral dipeptides are separated according to their hydrophobicities[4]. Apparently,

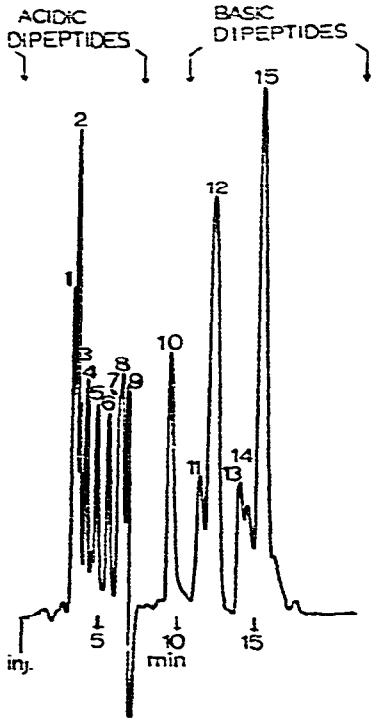

Fig. 1. HPLC separation of dipeptides with a mobile phase consisting of 0.038% (w/v) octanoic acid, 0.002% (w/v) pentanoic acid and 5% (v/v) butanol in water. 1 = Asp–Asp; 2 = Glu–Glu; 3 = Asp–Leu; 4 = Ala–Ala; 5 = Val–Ala; 6 = Glu–Leu; 7 = Ser–Ser; 8 = Val–Val; 9 = solvent disturbance; 10 = Arg–Asp; 11 = Lys–Ser; 12 = His–Ala; 13 = Arg–Val; 14 = Lys–Lys; 15 = His–Phe.

the enhancement in the capacity factors of basic dipeptides is due to the electrostatic interaction of the carboxylic anion and the protonated free amino groups of dipeptides. Due to electrostatic repulsion between the anion of the eluite and that of the carboxylic acid, acidic dipeptides are not retained. Whether the electrostatic interaction occurred in the mobile phase or at the modified surface charge brought about by partition of the carboxylic anion onto the stationary surface[17,18] is not clear.

The capacity factor of a dipeptide is a function of the chain length of the carboxylic acid. Fig. 2 shows the observations on acidic (panel A), neutral (panel B), dibasic (panel C) and monobasic dipeptides (panel D). Panel A indicates that as the chain length of carboxylic acid increases, the ln k' of acidic dipeptides including Glu–Leu, Asp–Leu, Leu–Phe, Glu–Glu and Asp–Asp decreases. It is not likely that the reduction of ln k' is solely the result of physiochemical changes of the mobile phase due to the ionization of carboxylic acids, because the dissociation constant of butyric acid (pK_a 4.81) is not significantly different from that of octanoic acid, pK_a 4.89. However, the hydrocarbonaceous stationary phase tends to adsorb more of the aliphatic carboxylic acids of longer chain length, resulting in increasing the surface charge of the stationary phase. The reduction in the retention of the acidic dipeptides may be due to the repulsive interaction of the surface charge with the anions of the acidic dipeptides.

Panel B of Fig. 2 shows the chain length dependency of ln k' of neutral dipeptides. The enhancement in ln k' *versus* carbon number of the carboxylic acids follows a sigmoidal curve. Using anion surfactants to increase the capacity factor of

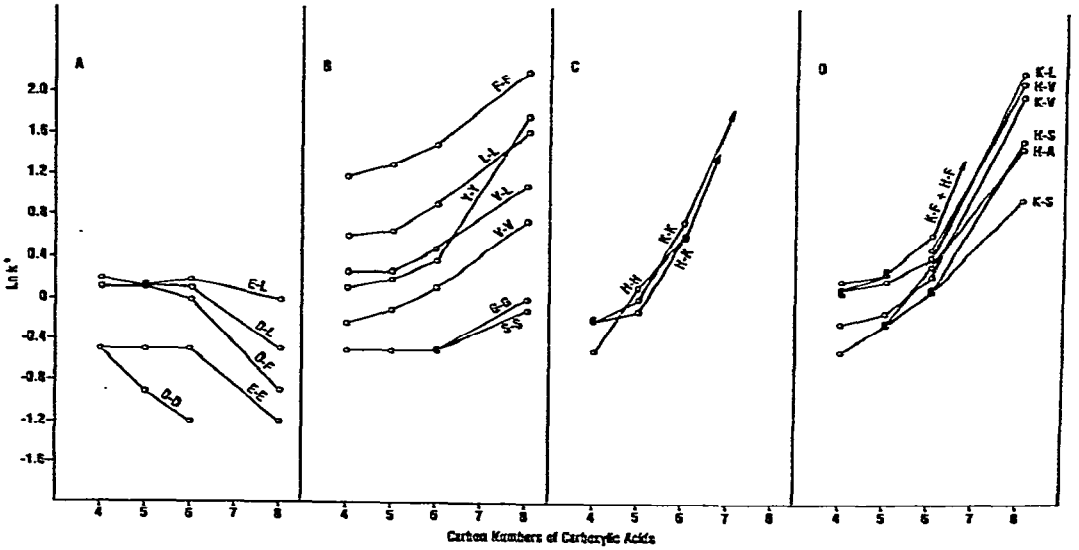

Fig. 2. The ln k' *vs.* chain length of carboxylic acid. The dipeptides include: (A) acidic; Glu–Leu (E–L), Asp–Leu (D–L), Asp–Phe (D–F), Glu–Glu (E–E), Asp–Asp (D–D); (B) neutral: Phe–Phe (F–F), Leu–Leu (L–L), Tyr–Tyr (Y–Y), Val–Leu (V–L), Val–Val (V–V), Gly–Gly (G–G), Ser–Ser (S–S); (C) dibasic: Lys–Lys (K–K), His–His (H–H), His–Lys (H–K); (D) monobasic: Lys–Phe (K–F), His–Phe (H–F), Lys–Leu (K–L), His–Val (H–V), Lys–Val (K–V), His–Ser (H–S), His–Ala (H–A), Lys–Ser (K–S). Mobile phase: 5% (v/v) butanol in water containing 0.03% (w/v) carboxylic acids with chain lengths of 4, 5, 6 and 8.

amino acids has been reported[4,19]. Panel B also reveals that the enhancement in the hydrophobicity of the eluites due to the carboxylic anion is additive to the existing hydrophobicity of the neutral dipeptides. Fig. 3 illustrates that if the total carbon number including the aliphatic side chains of the dipeptides Val–Val, Val–Leu, Leu–Leu and that of the carboxylic acids are plotted against the ln k', a linear relationship is observed. The slope of the diagram indicates that an increase of one methylene unit increase ln k' by 0.31.

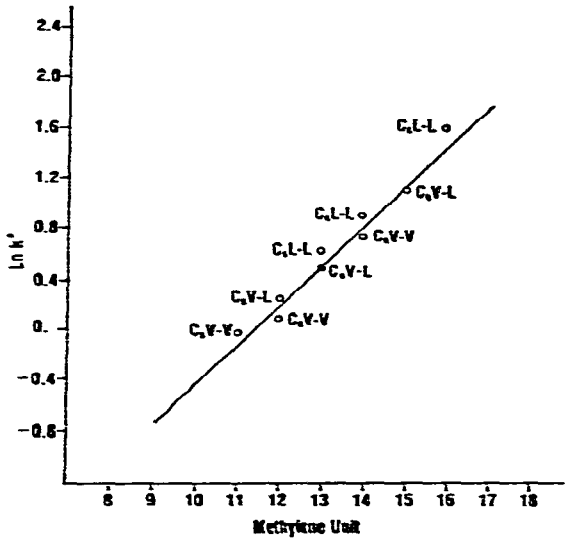

Fig. 3. The linear relationship between the natural logarithm of capacity factor for Val–Val (V–V), Val–Leu (V–L), and Leu–Leu (L–L) and the total carbon number including those of the aliphatic side chains of the dipeptides and that of the carboxylic acids. The composition of mobile phase is the same as that of Fig. 2 with C_5 being pentanoic acid; C_6, hexanoic acid; C_8, octanoic acid.

Panel C of Fig. 2 shows the chain length dependency of ln k' of dibasic dipeptides, including His–His, Lys–Lys and His–Lys. Since each compound contains three free amino groups, it was expected that carboxylic acid would have profound effects on the capacity factors. Indeed, adding 0.03% w/v octanoic acid in the mobile phase completely retains the dipeptides in the column. Furthermore, in contrast to panel B, the capacity factors of the three dipeptides are not much different, suggesting that enhancement in the hydrophobicity is mainly attributed to the aliphatic chain of the carboxylic acid. Panel D shows the chain length dependency of ln k' of monobasic dipeptides. Similar to panel B, ln k' as a function of chain length of carboxylic acid follows a sigmoidal curve, except with a steeper curvature concave upward. The observation is consistent because monobasic dipeptides contain one additional free amino group compared to neutral dipeptides, resulting in an additional basic site for electrostatic interaction with the carboxylic anion. It is interesting to note, as shown in Table I, that as the carbon number of the carboxylic acid is changed from six to eight, the ln k' of six monobasic dipeptides are increased by 1.6 ± 0.23 which is over twice as large as the increase in the ln k' of the four neutral dipeptides, 0.68 ± 0.03. From the

above observations, we conclude that by proper choice of the aliphatic carboxylic acid, both neutral and basic dipeptides can be separated from acidic dipeptides. The separation of dipeptides shown in Fig. 1 also suggests that it is possible that the effect of simultaneously using two carboxylic acids of different chain lengths would be of value in some cases.

TABLE I

CHANGES IN CAPACITY FACTORS OF DIPEPTIDES WITH INCREASE IN CARBON NUMBER OF CARBOXYLIC ACID

Dipeptides	*ln k'*		*Δln k'*
	C_6	C_8	
Monobasic			
His–Val	0.47	2.13	1.66
His–Ala	0.34	1.54	1.20
Lys–Phe	0.64	>2.22	1.58
Lys–Leu	0.41	2.21	1.80
Lys–Val	0.25	1.99	1.74
Mean ± S.D.			1.60 ± 0.23
Neutral			
Val–Val	0.10	0.74	0.64
Val–Leu	0.59	1.29	0.70
Leu–Leu	0.92	1.62	0.70
Phe–Phe	1.50	2.18	0.68
Mean ± S.D.			0.68 ± 0.03

Several publications[18,19] indicate that addition of inorganic salts to the mobile phase has drastic effects on the capacity factors of the eluites in the presence of surfactant ions. It is postulated that inorganic ions compete with the surfactant ion for the solute ion, resulting in a reduced capacity factor. Fig. 4 shows that the capacity factors of Leu–Leu, Val–Val fell to a constant value at a salt concentration of 7.5 mmole and remained constant by the addition of more salts. Since acidic dipeptides are not retained in this solvent system, the capacity factor of Asp–Asp as shown in Fig. 4 was not affected by the addition of salts. Increasing the organic solvent content in the mobile phase tends to reduce the capacity factor of the solutes[18]. The results shown in Fig. 5 indicate that increasing the butanol concentration in the eluent decreases the capacity factors of seven dipeptides.

Applications in polypeptide analysis

Over the past few years, this laboratory has developed a technique to analyze protein primary structure utilizing DAP I/IV to enzymatically cleave proteins into a mixture of dipeptides, followed by analysis of the dipeptides by GC–MS methods[13–15]. This method, while being fast and accurate, has some limitations, because basic dipeptides are not easily amenable to GC. In addition, when using the system at low nanomolar levels, there is a problem of interference in the GC–MS dipeptide analysis due to the presence of the proteases and also low-molecular-weight constituents found in the hydrolysates. HPLC seemed capable of solving these two problems especially in

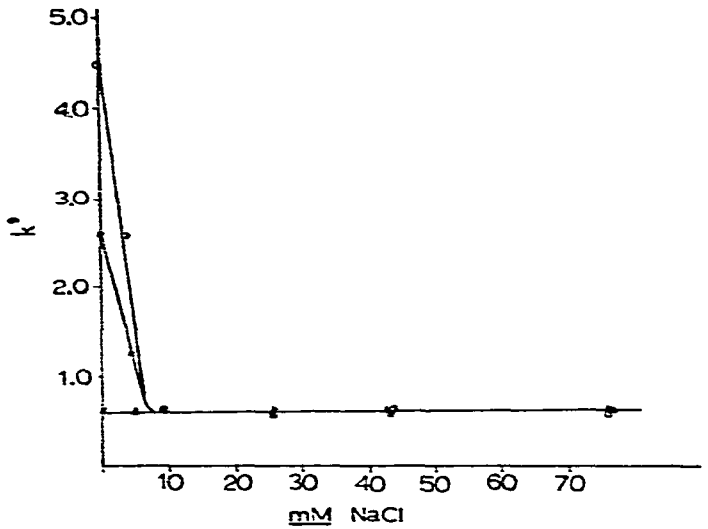

Fig. 4. Salt effect on the capacity factors of dipeptides: ○, Lys–Lys; ×, Val–Val and △, Asp–Asp. The mobile phase consisted of 20% (v/v) acetonitrile and 0.03% (w/v) octanoic acid in water.

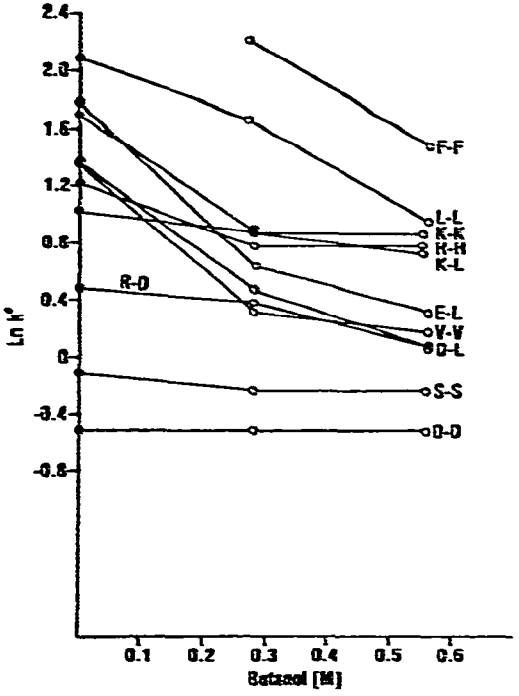

Fig. 5. Effect of butanol concentration on the capacity factors of dipeptides: Phe–Phe (F–F); Leu–Leu (L–L); Lys–Lys (K–K); His–His (H–H); Lys–Leu (K–L); Glu–Leu (E–L); Val–Val (V–V); Arg–Asp (R–D); Asp–Leu (D–L); Ser–Ser (S–S), and Asp–Asp (D–D). The mobile phase consisted of 0.032% (v/v) pentanoic acid in water.

view of the results just described. Lyophilized hydrolysates could be taken up in water and injected directly into the HPLC. With an aid of a pre-column, the proteases and any substances larger than approximately 10,000 molecular weight could be readily removed. Separation of basic dipeptides from acidic and neutral dipeptides also facilitated their identifications by normal chromatographic procedures or by collecting the suspected peaks from HPLC for further characterization.

Val-4-angiotensin III with known amino acid sequence as Arg–Val–Tyr–Val–His–Pro–Phe was enzymatically digested according to the procedures described before. It is worth noting that the dipeptide containing histidine (His–Pro) and arginine (Arg–Val), need special chemical treatment in order to obtain reliable structural information by GC–MS techniques. The guanidine moiety of arginine needs to be protected prior to the formation of pentafluoropropionyl dipeptide methyl ester[13,16]. The lability of the acetylated imidazole ring causes rapid deacylation of the histidine side chain, especially if another polyfunctional residue is present in the dipeptide. Panel A in Fig. 6 shows the separation of the three authentic dipeptides Arg–Val, Tyr–Val and His–Pro which were anticipated to be relased from val-4-angiotensin III. Panel B shows the HPLC of an incubation mixture containing no substrate. The dipeptide profile resulting from enzymatic digestion is shown in panel C. Panel D provides tentative confirmation of the dipeptides by coelution of the incubation mixture with the solution of panel A. No further confirmation was performed.

Another polypeptide, ribonuclease S, consisting of twenty amino acid residues: Lys–Glu–Thr–Ala–Ala–Ala–Lys–Phe–Glu–Arg–Gln–His–Met–Asp–Ser–Ser–Thr–Ser–Ala–Ala was used as substrate. Nine different dipeptides would be anticipated

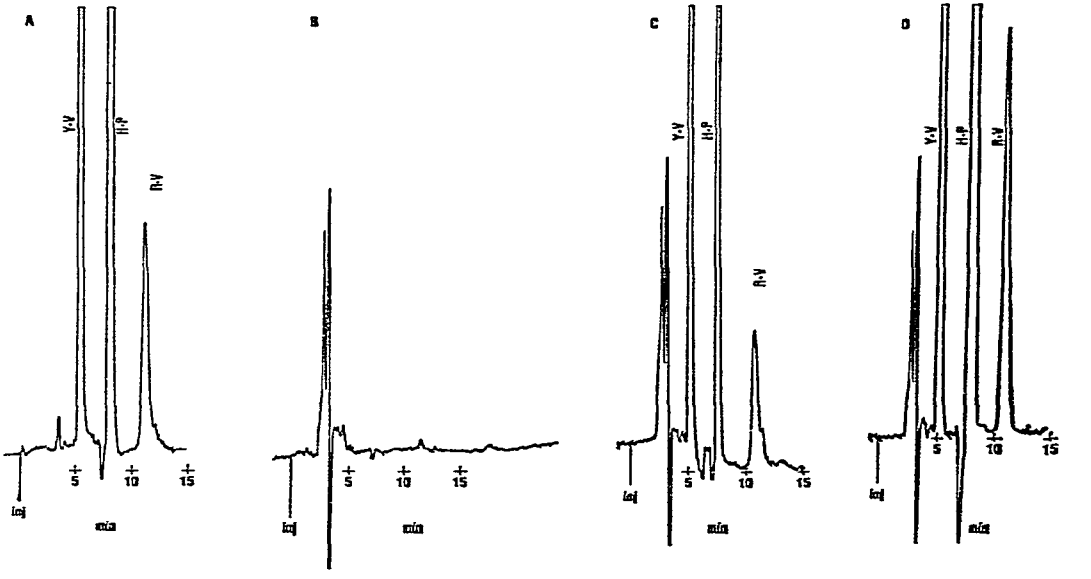

Fig. 6. HPLC separation of DAP I/IV hydrolysates of val-4-angiotensin III; (A), separation of authentic Arg–Val (R–V), Tyr–Val (Y–V) and His–Pro (H–P); (B), HPLC of an incubation mixture containing no substrate; (C), dipeptide profile of DAP I/IV hydrolysates; (D), confirmation of the dipeptides by co-injection of solutions A and C.

in the incubation mixture, with Lys–Glu, Lys–Phe, Gln–His and Gly–Arg as basic dipeptides. Chromatography of these dipeptides by HPLC is shown in Fig. 7. While it was not possible to determine the acidic and neutral dipeptides in the hydrolysate based on Fig. 7, their total collection and analysis by GC/MS was easily done. HPLC eluent containing total dipeptides was collected and lyophilized. This residue could be

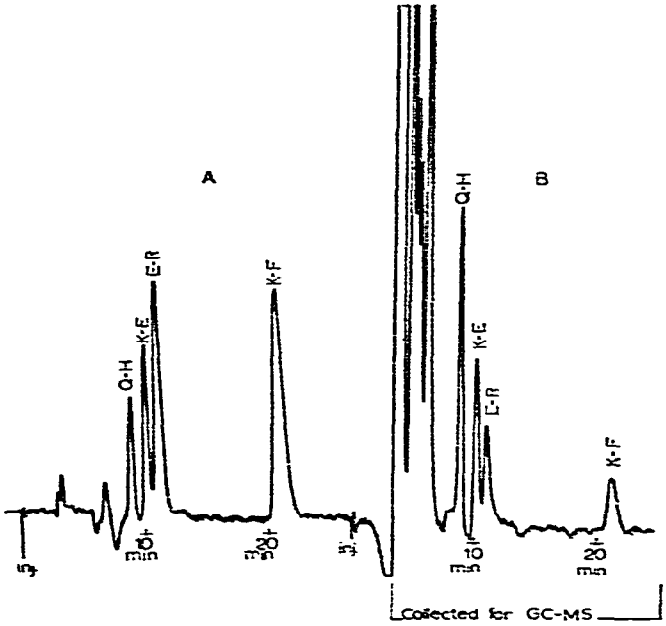

Fig. 7. HPLC separation of DAP I/IV hydrolysates of ribonuclease S peptide. The mobile phase composed of 0.03% (w/v) octanoic acid in 5% (v/v) butanol in water. (A), Separation of authentic Gln–His (Q–H), Lys–Glu (K–E), Glu–Arg (E–R) and Lys–Phe (K–F); (B), Dipeptide profile of DAP I/IV hydrolysates.

derivatized for GC–MS analysis without need of further purification. The upper panel of Fig. 8 shows the total ion chromatogram of dipeptides in the HPLC eluent. Each derivatized dipeptide designated in Fig. 8 was confirmed by comparing its mass spectrum to that of the authentic compound. As a comparison, the lower panel of Fig. 8 shows the total ion chromatogram from GC–MS analysis of the same incubation mixture but without HPLC purification. As anticipated, many other peaks appear in the chromatogram which overlap and obscure those of the dipeptides, especially Ala–Ala, Ser–Ser, Met–Asp and Lys–Phe. The GC–MS analysis did not show the dipeptides Gln–His and Glu–Arg for the same reasons as described before. However, the proper HPLC fraction can be collected for further characterization by alternative techniques. As an example, the fraction corresponding to Glu–Arg was collected; taken to dryness, and the residue treated according to the procedures described for arginine containing dipeptides. The final product was analyzed by solid probe MS. The mass spectrum, shown in Fig. 9, is identical to that of PFP-glutamyl (4,6-dimethyl pyrimid-2-yl)-ornithine dimethyl ester, the product of Glu–Arg after derivatization.

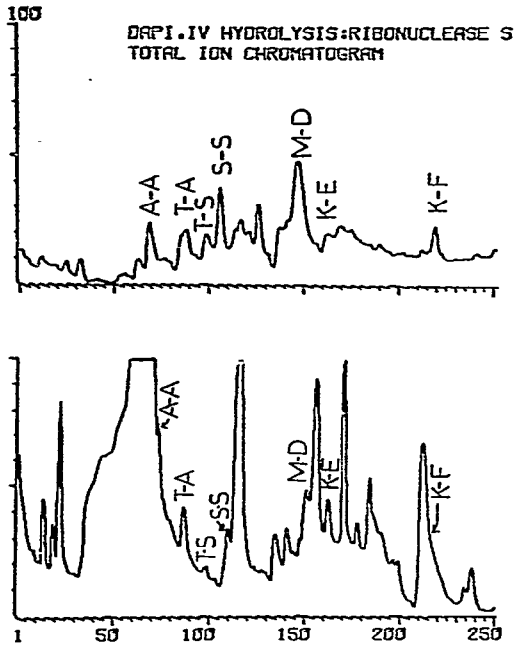

Fig. 8. Total ion chromatograms of dipeptides in DAP I/IV hydrolysates of ribonuclease S-peptide by GC–MS technique. Upper panel, after HPLC purification; lower panel, without HPLC purification. The dipeptides were identified by comparing their mass spectra with those of the authentic compounds: Ala–Ala (A–A), Thr–Ala (T–A), Thr–Ser (T–S), Ser–Ser (S–S), Met–Asp (M–D), Lys–Glu (K–E), Lys–Phe (K–F).

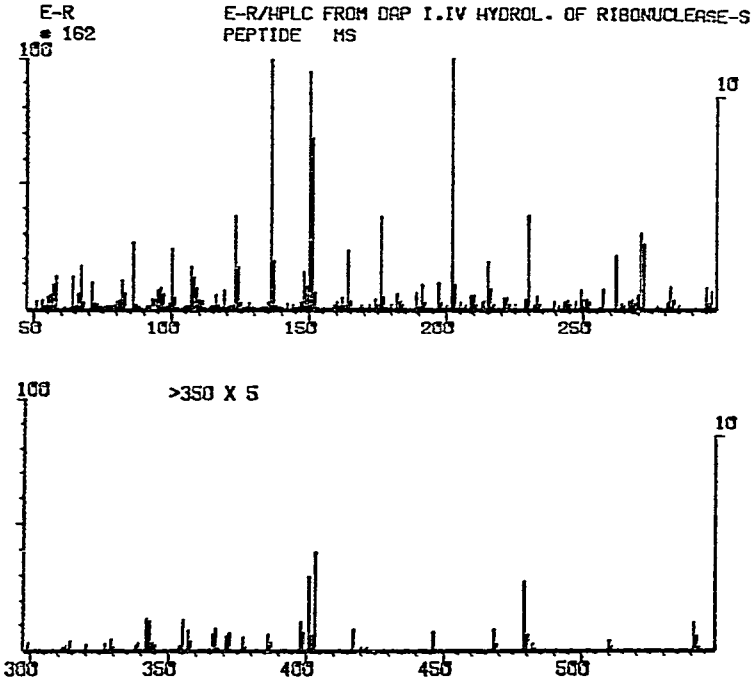

Fig. 9. Mass spectrum of PFP-glutamyl (4,6-dimethylpyrimidyl)-ornithine dimethyl ester from DAP I/IV digestion of ribonuclease S-peptide after HPLC purification. *m/z* 541 is the molecular ion.

In conclusion, by adding small amounts of aliphatic carboxylic acid as an additive to the mobile phase, reversed-phase HPLC was shown to be a useful technique to compliment the GC–MS techniques, especially in regard to the analysis of basic dipeptides. The sigmoidal curves of the dependence of ln k' of basic dipeptides on chain length of carboxylic acids indicate that even the most basic dipeptides such as His–His, Lys–Lys and Arg–Arg, can be properly separated by manipulation of different carboxylic acids or their combinations as ion pair reagents. Furthermore, the HPLC method described herein provides a fast and efficient means to purify dipeptides from biologic samples for futher identification by alternate techniques.

ACKNOWLEDGEMENTS

We thank the Robert A. Welch Foundation (Grant AU-671) and the National Institutes of Health (Grant No. 1 ROL GM 25405) for their generous support.

REFERENCES

1 D. H. Spackman and W. H. Stein, *J. Biol. Chem.*, 235 (1960) 648.
2 S. B. H. Kent, A. R. Mitchell, G. Barany and R. B. Merrifield, *Anal. Chem.*, 50 (1978) 155.
3 J. M. Manning and S. Moore, *J. Biol. Chem.*, 243 (1968) 5591.
4 I. Molnár and Cs. Horváth, *J. Chromatogr.*, 142 (1977) 623.
5 K. Krummen and R. W. Frei, *J. Chromatogr.*, 132 (1977) 27.
6 E. P. Kroeff and D. J. Pietrzyk, *Anal. Chem.*, 50 (1978) 1353.
7 D. J. Pietrzyk, E. P. Kroeff and T. D. Rotsch, *Anal. Chem.*, 50 (1978) 497.
8 E. P. Kroeff and D. J. Pietrzyk, *Anal. Chem.*, 50 (1978) 502.
9 W. Mönch and G. Dehnen, *J. Chromatogr.*, 147 (1978) 415.
10 J. A. Feldman, M. L. Cohn and D. Blair, *J. Liquid Chromatogr.*, 1 (1978) 833.
11 J. H. Knox and G. R. Laird, *J. Chromatogr.*, 122 (1976) 17.
12 B. Fransson, K.-G. Wahlund, I. M. Johansson and G. Schill, *J. Chromatogr.*, 125 (1976) 327.
13 W. E. Seifert, Jr., R. E. McKee, C. F. Beckner and R. M. Caprioli, *Anal. Biochem.*, 88 (1978) 149.
14 W. E. Seifert, Jr. and R. M. Caprioli, *Biochemistry*, 17 (1978) 436.
15 R. M. Caprioli and W. E. Seifert, Jr., *Biochem. Biophys. Res. Commun.*, 64 (1975) 295.
16 V. Bacon, E. Jellum, W. Patton, W. Pereira and B. Halpern, *Biochem. Biophys. Res. Commun.*, 37 (1969) 878.
17 P. T. Kissinger, *Anal. Chem.*, 49 (1977) 883.
18 Cs. Horváth, W. Melander, I. Molnár and P. Molnár, *Anal. Chem.*, 49 (1977) 2295.
19 J. C. Kraak, K. M. Jonker and J. F. K. Huber, *J. Chromatogr.*, 142 (1977) 671.