HYDROLYSIS OF POLYPEPTIDES AND PROTEINS UTILIZING A MIXTURE OF DIPEPTIDYLAMINOPEPTIDASES WITH ANALYSIS BY GC/MS.

R. M. Caprioli and W. E. Seifert†
Department of Chemistry, Purdue University
West Lafayette, Indiana 47907

Received March 11,1975

Summary

The hydrolysis of polypeptides and proteins to dipeptides using a mixture of dipeptidylaminopeptidase I and IV is demonstrated for several compounds. The dipeptide products were identified by gas chromatography/mass spectrometry of their acyl, ester derivatives. Partial hydrolysis of carboxymethylated soybean trypsin inhibitor (mw 20,327) shows the viability of the methods with an intact protein. Polypeptides with N-terminal arginyl or lysyl residues, which cannot be hydrolyzed by dipeptidylaminopeptidase I alone, are readily hydrolyzed by the dipeptidylaminopeptidase I/IV mixture, as shown by the digestion of ribonuclease S-peptide. The enzyme mixture releases dipeptides from the N-terminus of a polypeptide or protein with hydrolysis being terminated only when proline appears in position three of a remaining polypeptide fragment. The techniques are discussed in terms of new types of analyses which can be performed using this combined enzyme-GC/MS method.

Dipeptidylaminopeptidases are a class of proteolytic enzymes found in a variety of mammalian tissue which catalyze the hydrolysis of polypeptides to dipeptides starting from the N-terminus of the polypeptide substrate. Dipeptidylaminopeptidase I (DAP I)* was first described by Gutman and Fruton[1] and was later more extensively studied by McDonald and coworkers.[2] The specificity of DAP I is somewhat limited in that it cannot hydrolyze a polypeptide when the peptide bond to be cleaved involves a proline residue, or if the polypeptide chain contains an N-terminal arginine or lysine residue. Other peptide bonds are cleaved at varying rates. The possibility of employing DAP I for the determination of amino acid sequences of polypeptides was

†David Ross Fellow, Purdue Research Foundation. Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, Purdue University.

*The abbreviations used in the text are: DAP for dipeptidylaminopeptidase, GC/MS for gas chromatography/mass spectrometry, and PFP for pentafluoropropionyl.

suggested by McDonald et al.[3] in 1969, and subsequent work by Callahan et al.[4] demonstrated the use of several physical techniques for the identification of the dipeptides, including paper and ion-exchange chromatography, and membrane diffusion techniques.

Dipeptidylaminopeptidase IV (DAP IV) has been isolated from porcine kidney as well as other sources by Hopsu-Havu and coworkers[5] and has been shown to hydrolyze dipeptides from the N-terminus of peptides when the penultimate residue is proline, liberating the dipeptide "x-pro". However, this enzyme will not cleave the imide bond of proline when this residue lies in position three of a peptide chain. No work has yet been published concerning its action with other than small peptides containing 3-4 residues.

The key in using these enzymes for the analysis of polypeptides and proteins lies in the employment of a method of identification of the dipeptides which is fast, accurate, and can be used routinely on 50-100 nanomoles of sample. Recent work[6,7] has shown that gas chromatography/mass spectrometry (GC/MS) offers such advantages.

In the present paper, the usefulness of the combined DAP - GC/MS procedure is demonstrated by the hydrolysis and analysis of several polypeptides and partial hydrolysis of a protein. DAP I and DAP IV are used in mixtures to eliminate most of the difficulties with enzyme specificity when either enzyme is used alone. The utility of the technique for several types of analyses of polypeptides and proteins, including sequencing, is discussed.

Materials and Methods

DAP I was isolated from bovine spleen and was assayed and stored according to McDonald et al.[8] The activity of the preparation was approximately 12 U/mg. DAP IV was isolated from porcine kidney according to Hopsu-Havu et al.[5]. The specific activity was found to be approximately 9 U/mg.

The hydrolysis of polypeptides and proteins was carried out at pH 6.5 in the following manner. The peptide was initially dissolved in 5% 2,6-dimethylpyridine (30 μl/0.1 μmole) and diluted with water (79 μl/0.1 μmole). To this

solution was added 0.5% acetic acid (40 µl/0.1 µmole), 0.1 N HCl (47 µl/0.1 µmole), 0.375 M 2-mercaptoethanol (8 µl/0.1 µmole), and 0.1 M EDTA (2 µl/0.1 µmole). This solution was equilibrated at 37° for five minutes and the hydrolysis initiated by adding 3-5 units of DAP I and 2-3 units DAP IV per micromole of peptide. Digestion was carried out for 2-3 hours and then stopped by freezing and lyophilization.

The residue from the enzyme hydrolysis was treated in the following manner to convert the dipeptides to the N,O-perfluoropropionyl (PFP), methyl ester derivatives for GC/MS analysis. To the lyophilized residue in a 1 ml Reactivial was added 200 µl of dry methanol, the suspension cooled in a dry ice-ethanol bath and 40 µl of thionyl chloride slowly added. The solution was allowed to come to room temperature and then heated in a constant temperature block at 45° for 20 min. The reagents were removed in vacuo and 100 µl of PFP anhydride (Pierce Chemical Co.) was added. After 15 min at room temperature, the reagents were again removed in vacuo and the residue dissolved in 10-50 µl of dry dioxane. After thorough mixing, 1-2 µl of this solution was injected into the GC inlet of the GC/MS. Under these derivatization conditions, the side chain amides are converted to their acids. When analyses are performed to distinguish the acids and amides, the same derivatization procedure can be used except the thionyl chloride/methanol mixture must be maintained at 0°. Also, the guanido group of arginine is not derivatized by either procedure and thus dipeptides containing this residue will not emerge from the GC. To analyze these, the hydrolytic mixture is first treated with 2,4-pentanedione to convert arginyl residues to their corresponding 2,4-dimethylpyrimidyl analogs, as previously described.[9] The derivatives are then treated and analyzed as above.

The Edman degradation were performed on ribonuclease S-peptide according to the procedure of Blombäck et al.[10]

A Finnigan 3200/6000 GC/MS/data system was employed for the analyses of the dipeptide products. Separation was effected using a 0.2 x 45 cm column

packed with 3% Dexsil 300 on 100/120 mesh chromosorb G (Applied Science Labs). On-column injection was used with the injector at 220° and a linear temperature program from 100° to 250° at 10°/min. Helium was used as the carrier gas at a flow of approximately 30 ml/min. Mass spectra were obtained using an electron-impact source operated at 70 eV and at a temperature of 100°. Ion chromatograms, plots analogous to normal GC traces, were obtained by plotting total ion intensity for each mass spectrum taken in the analysis versus mass spectrum number.

S-Carboxymethylated soybean trypsin inhibitor (Kunitz) was generously supplied by Dr. Peter Fankhauser, Department of Chemistry, Purdue University. Ribonuclease S-peptide was obtained from the Sigma Chemical Co.

Results and Discussion

DAP I and DAP IV have nearly complementary specificities. DAP I generally hydrolyzes all peptide bonds except those involving proline and those where the N-terminus is arginine or lysine. DAP IV preferentially hydrolyzes a prolyl containing polypeptide when this residue is in the penultimate position and present work shows that it releases dipeptides from a polypeptide having N-terminal lysine or arginine. A mixture of DAP I and DAP IV should therefore release dipeptides from a polypeptide chain and should terminate hydrolysis only when the remaining polypeptide fragment has a prolyl residue in position three. This is illustrated in Figure 1 for the partial hydrolysis of S-carboxymethylated soybean trypsin inhibitor, a protein containing 181 residues. The sequence of the N-terminal portion of this protein is;

asp-phe-val-leu-asp-asn-glu-gly-asn-pro-leu-glu-asn-gly-gly-thr-
tyr-tyr-ile-leu-ser-asp-ile-thr-ala-phe-gly-gly-ile-arg-ala-ala-pro---

Proline appears in positions 10 and 33. Since the first is an even numbered position, the DAP I/IV mixture hydrolyzes through this residue, eventually stopping after hydrolyzing the peptide bond between positions 30 and 31. This places proline in position 3 of the remaining polypeptide core. Thus, fifteen dipeptides were liberated in the hydrolysis, fourteen of which are shown in Figure 1. Since in this analysis arginine was not specifically deri-

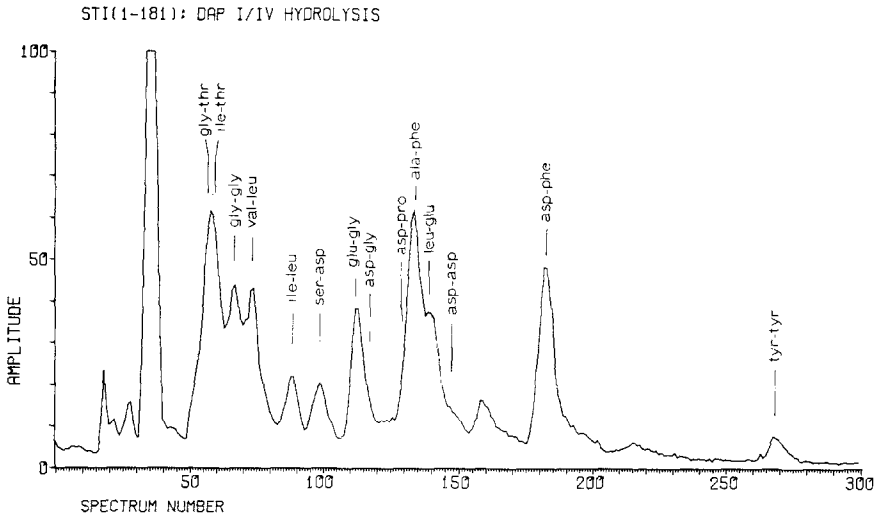

Figure 1. Ion chromatogram of dipeptides liberated from S-carboxymethylated soybean trypsin inhibitor by digestion with the DAP I/IV mixture. The dipeptides were derivatized as described in the text.

vatized, the dipeptide ile-arg did not emerge from the GC. The mass spectra of two of these, asp-phe and val-leu, are given in Figure 2. Also, under the derivatization conditions employed, asparagine and glutamine were converted to aspartic and glutamic acids, respectively, as seen in Figure 1. A second analysis was performed to analyze the amides, as described in the experimental section, and the following dipeptides were identified; asp-asn, asn-pro, and asn-gly. Although this low temperature derivatization is not the method of choice because of less quantitative derivatization, it does permit identification of the amides when necessary.

The hydrolysis of soybean trypsin inhibitor illustrates several important points about the methods employed. First, DAP I and DAP IV are compatible with one another and can be used in a mixture for the hydrolysis of a substrate. At pH 6.5 both have sufficient activity to perform the hydrolysis and do not interfere with one another. Second, although complete hydrolysis of the intact protein was not effected due to proline in position 33, this work demonstrates the viability of the methods for the analysis of molecules of the size of

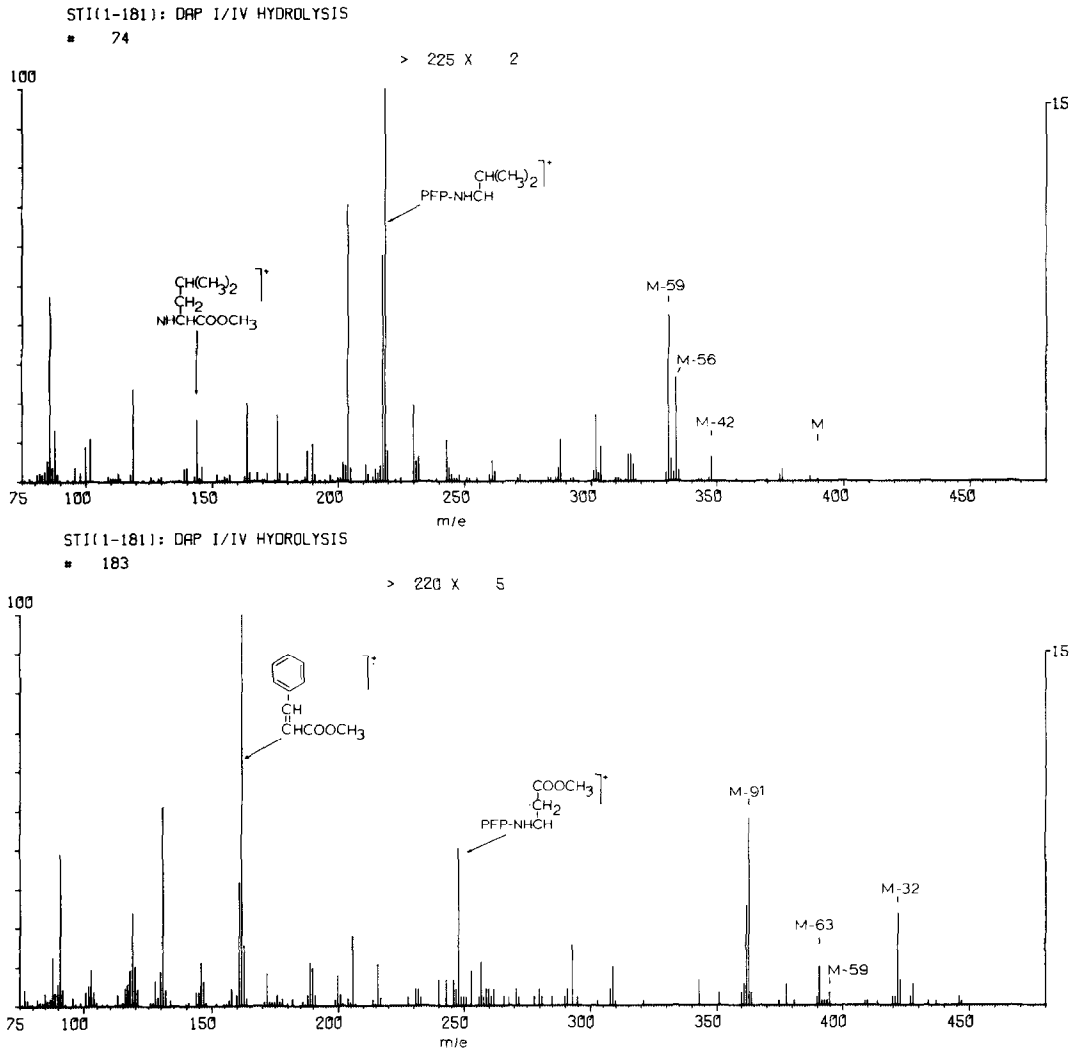

Figure 2. Mass spectra of derivatives of two of the dipeptides formed from digestion of S-carboxylmethylated soybean trypsin inhibitor; top, N-PFP-val-leu methyl ester; bottom N-PFP-asp-phe dimethyl ester.

proteins. This is indeed significant since it may obviate the need to perform tryptic, chymotryptic, or other partial hydrolyses to form smaller fragments and the time-consuming separation of the various fragments which must follow. Finally, it shows that the enzyme mixture can in fact hydrolyze past prolyl residues if these lie in even-numbered positions in the polypeptide chain.

The use of DAP IV allows polypeptides to be hydrolyzed which contain lysine

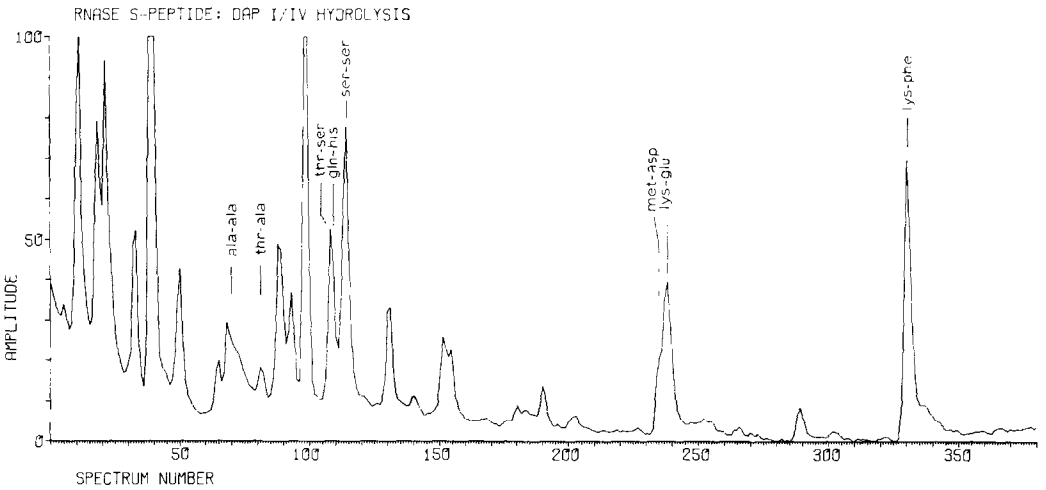

Figure 3. Ion chromatogram of dipeptides liberated by the digestion of ribonuclease S-peptide with the DAP I/IV mixture. The dipeptides were derivatized as described in the text.

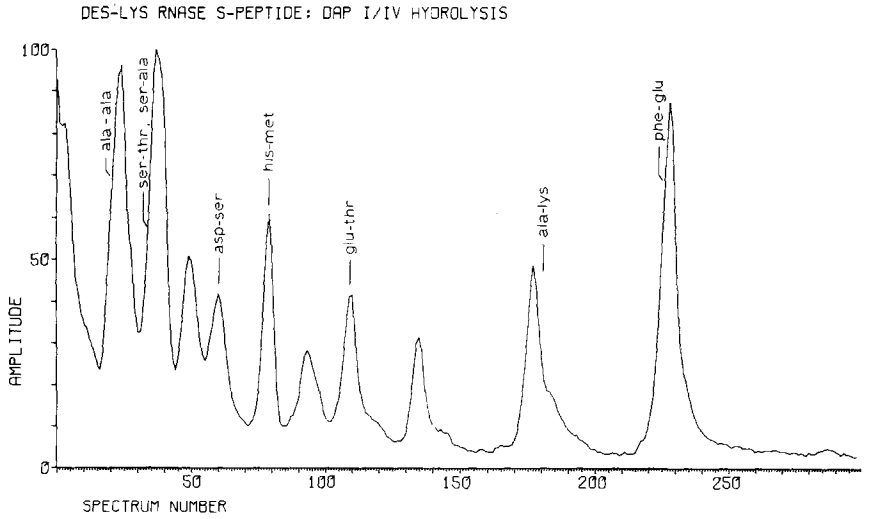

Figure 4. Ion chromatogram of dipeptides liberated by the digestion of des-lys ribonuclease S-peptide with the DAP I/IV mixture. The dipeptides were derivatized as described in the text.

---

and arginine in the N-terminus. Figure 3 shows the ion chromatogram obtained from the analysis of dipeptides liberated from ribonuclease S-peptide. This polypeptide contains 20 residues and has the sequence;

lys-glu-thr-ala-ala-ala-lys-phe-glu-arg-gln-his-met-asp-ser-ser-thr-ser-ala-ala

The dipeptides giving rise to the peaks at spectrum numbers 238 and 330 of Figure 3 are lys-glu and lys-phe, respectively. These together with the various other dipeptides demonstrate that the DAP I/IV mixture can hydrolyze past N-terminal lysyl residues. Similarly, Figure 4 shows the ion chromatogram for des-lys ribonuclease S-peptide prepared by performing one round of the Edman reaction on the S-peptide. This results in the arginyl residue appearing in the N-terminal position after partial hydrolysis by DAP I/IV. As seen in Figure 4, dipeptides occurring after arg-gln were also released, i.e., his-met, asp-ser, ser-thr and ser-ala, demonstrating that the DAP I/IV mixture hydrolyzed past the N-terminal arginyl residue.

In conclusion, the work presented above shows the usefulness of the DAP I/IV mixture in hydrolyzing polypeptides. These exoproteases provide another dimension for the specific hydrolysis of polypeptides, terminating when a prolyl residue resides in position three of a polypeptide fragment. Dipeptides containing N-terminal arginyl and lysyl residues and even-positioned prolyl residues, stopping points when DAP I is used alone, are released by DAP IV. The enzyme mixture is able to hydrolyze proteins of molecular weight of approximately 20,000, and possibly greater. This work permits several different types of protein analyses to be effected; i) for unknowns, the primary sequence may be determined by hydrolysis and analysis of the original polypeptide and the des-N-terminal residue polypeptide, as previously described,[4,7] and ii) for partially or completely known sequences, the position of modified residues or those residues which may have been incompletely or incorrectly identified by other procedures may be determined. With just a single hydrolysis of a polypeptide chain, identification of the dipeptides would provide a type of nearest neighbor analysis for that sample. The use of this method for sequence checking is particularly appealing since the method is fast, uses relatively small amounts of sample and relies on chemistry and analytical techniques different from those used in more classical procedures. Thus, it can be expected to have unique strengths and weaknesses.

References

1. Gutman, H. R., and Fruton, J. S. (1948) J. Biol. Chem. 174, 851-858.
2. McDonald, J. K., Ellis, S., and Reilly, T. J., (1966) J. Biol. Chem. 241, 1494-1501.
3. McDonald, J. K., Zeitman, B. B., Reilly, T. J., and Ellis, S. (1969) J. Biol. Chem. 244, 2693-2709.
4. Callahan, P. K., McDonald, J. K., and Ellis, S. (1972) in Methods of Enzymology, XXV, (C. H. W. Hirs and S. N. Timasheff, eds.), pp. 282-298, Academic Press, New York.
5. Hopsu-Havu, V. K., Rintola, P., and Glenner, G. G. (1968) Acta Chem. Scand. 22, 299-308.
6. Ovchinnikov, Y. A., and Kiryushkin, A. A. (1972) FEBS Lett. 21, 300-302.
7. Caprioli, R. M., Seifert, W. E., and Sutherland, D. E. (1973) Biochem. Biophys. Res. Comm. 55, 67-75.
8. McDonald, J. K., Callahan, P. X., and Ellis, S. (1972) in Methods of Enzymology, XXV, (C. H. W. Hirs and S. N. Timasheff, eds.), pp. 272-281, Academic Press, New York.
9. Morris, H. R., Dickinson, R. J., and Williams, D. H. (1973) Biochem. Biophys. Res. Commun. 51, 247-255.
10. Blömback, B., Blömback, M., Edman, P., and Hessel, B. (1966) Biochim. Biophys. Acta 115, 371-396.