study on the purified enzyme, the loss of both enzymic activity and coenzyme binding by derivatization with 1 equiv of MalNEt. Although several soluble dehydrogenases have essential sulfhydryl groups, modification of these residues did not prevent coenzyme binding as in the present study [liver alcohol dehydrogenase (Li & Vallee, 1963); yeast alcohol dehydrogenase (Dickinson, 1972); lactate dehydrogenase (Holbrook, 1966)]. Our results suggest that D-β-hydroxybutyrate dehydrogenase is unique, compared to these soluble dehydrogenases, in that both phospholipid and a sulfhydryl moiety on the enzyme are required for coenzyme binding.

References

Blair, P. V. (1967) *Methods Enzymol. 10*, 78–81.

Bock, H.-G. O., & Fleischer, S. (1974) *Methods Enzymol. 32*, 374–391.

Bock, H.-G. O., & Fleischer, S. (1975) *J. Biol. Chem. 250*, 5774–5781.

Brewer, J. M., Pesce, A. J., & Ashworth, R. B. (1974) *Experimental Techniques in Biochemistry*, pp 3–4, Prentice-Hall, Englewood Cliffs, NJ.

Chen, P. S., Toribara, T., & Warner, H. (1956) *Anal. Chem. 28*, 1756–1758.

Dickinson, F. M. (1972) *Biochem. J. 126*, 133.

Fleischer, S., & Klouwen, H. (1961) *Biochem. Biophys. Res. Commun. 5*, 378–393.

Fleischer, S., Rouser, G., Fleischer, B., Casu, A., & Kritschevsky, G. (1967) *J. Lipid Res. 8*, 170–180.

Fleischer, S., Bock, H.-G. O., & Gazzotti, P. (1974) *Membr. Proteins Transp. Phosphorylation, Proc. Int. Symp.*, 125–136.

Fleischer, S., McIntyre, J. O., Stoffel, W., & Tunggal, B. D. (1979) *Biochemistry 18*, 2420–2429.

Fluharty, A., & Sanadi, D. R. (1961) *J. Biol. Chem. 236*, 2772.

Gazzotti, P., Bock, H.-G. O., & Fleischer, S. (1974) *Biochem. Biophys. Res. Commun. 58*, 309–315.

Gazzotti, P., Bock, H.-G. O., & Fleischer, S. (1975) *J. Biol. Chem. 250*, 5782–5790.

Holbrook, J. J. (1966) *Biochem. Z. 344*, 141.

Isaacson, Y. A., Deroo, P. W., Rosenthal, A. F., Bittman, R., McIntyre, J. O., Bock, H.-G. O., Gazzotti, P., & Fleischer, S. (1979) *J. Biol. Chem. 254*, 117–126.

Latruffe, N., & Gaudemer, Y. (1974) *Biochimie 56*, 435–444.

Latruffe, N., Brenner, S. C., & Fleischer, S. (1979) *Abstr., Int. Cong. Biochem., 11th*, 270.

Li, T.-K., & Vallee, B. L. (1963) *Biochem. Biophys. Res. Commun. 12*, 44.

Lowry, O. H., Rosebrough, N. J., Farr, A. L., & Randall, R. J. (1951) *J. Biol. Chem. 193*, 265–275.

Nielsen, N. C., & Fleischer, S. (1973) *J. Biol. Chem. 248*, 2549–2555.

Nielsen, N. C., Zahler, W. L., & Fleischer, S. (1973) *J. Biol. Chem. 248*, 2556–2562.

Riordan, J. F., & Vallee, B. L. (1967) *Methods Enzymol. 11*, 541–548.

Ross, E., & Schatz, G. (1973) *Anal. Biochem. 32*, 91–100.

Rouser, G., & Fleischer, S. (1967) *Methods Enzymol. 10*, 385–406.

Sekuzu, I., Jurtshuk, P., Jr., & Green, D. E. (1963) *J. Biol. Chem. 238*, 975–982.

Tan, A. W. H., Smith, C. M., Aogaichi, T., & Plaut, G. W. E. (1975) *Arch. Biochem. Biophys. 166*, 164–173.

Velick, S. F. (1958) *J. Biol. Chem. 233*, 1455–1467.

Wise, J. B., & Lehninger, A. L. (1962) *J. Biol. Chem. 237*, 1363–1370.

Determination of Polypeptide Amino Acid Sequences from the Carboxyl Terminus Using Angiotensin I Converting Enzyme†

Henry C. Krutzsch

ABSTRACT: A method for sequence analysis of polypeptides starting at the carboxyl terminus is described that utilizes degradation of the polypeptide into dipeptides with angiotensin I converting enzyme. Dipeptides were identified by gas chromatography–mass spectroscopy. Dipeptide alignment was achieved by replicate digestion of the polypeptide after modification at the carboxyl terminus either by chemical or enzymatic removal of one residue or by addition of a single residue. The addition reaction involved coupling of L-α-aminobutyric acid under conditions described herein which yielded essentially complete conversions. Unlike sequence determination methods that commence from the polypeptide amino terminus, this procedure does not require that a polypeptide have a free amino terminus for successful application. A number of polypeptides with varying chain lengths (up to 49 residues), containing among them most of the common amino acids, have been successfully analyzed in amounts as low as 5 nmol.

Most of the widely applied and well-established methods for the determination of polypeptide primary structure are based on sequential degradations that start from the amino terminus. Although these procedures most frequently utilize variations of the Edman degradation (Edman, 1970), they may also be accomplished by the rapid and sensitive dipeptidyl aminopeptidase (DAP)[1] method (Krutzsch & Pisano, 1977; Seifert & Caprioli, 1978). However, amino-terminal sequence methods fail completely if a polypeptide has a blocked N terminus or give only partial results if the sequence deter-

† From the Laboratory of Immunogenetics, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20205. *Received February 21, 1980.*

[1] Abbreviations used: DAP, dipeptidyl aminopeptidase; C, carboxy; N, amino; DCP, dipeptidyl carboxypeptidase; GC–MS, gas chromatography–mass spectrometry; ACE, angiotensin I converting enzyme; NMM·HOAc, *N*-methylmorpholine acetate; FA-F-G-G, furanacryloyl-L-phenylalanylglycylglycine; Me_3Si, trimethylsilyl; BSTFA, *N,O*-bis(trimethylsilyl)trifluoroacetamide; DMAC, *N,N*-dimethylacetamide; Abu-O-*t*-Bu, L-α-aminobutyric acid *tert*-butyl ester; EEDQ, *N*-(ethoxycarbonyl)-2-ethoxy-1,2-dihydroquinoline; CPB, carboxypeptidase B; CDI, 1,1′-carbonyldiimidazole; Tris, 2-amino-2-(hydroxymethyl)-1,3-propanediol.

mination ends before reaching the C terminus of the polypeptide. As a means of overcoming these problems, several chemical (Stark, 1972) and enzymatic (Ambler, 1972) procedures for polypeptide sequence determination have been outlined which start from the polypeptide carboxyl terminus. However, these methods have not been routinely applied for analysis of more than 3–5 residues of the polypeptide sequence. The limitations inherent in these methods stem from either the chemical side reactions that terminate C-terminal single-step degradations prematurely or the uncertainty of residue placement when carboxymonopeptidase enzymes are used to determine the sequence of more than 3–5 C-terminal residues.

The development of an improved method for extensive polypeptide sequence analysis that began at the C terminus would thus significantly aid and extend existing techniques for polypeptide structural studies. One possible technique for this type of polypeptide sequence determination might involve use of dipeptidyl carboxypeptidase (DCP) enzymes to degrade polypeptides into dipeptide fragments in a sequential manner from the C terminus. Possible enzymes for this use include angiotensin-converting enzyme (Holmquist et al., 1979) and the DCP enzyme isolated from *Escherichia coli* (Yaron, 1976). Rapid and sensitive techniques for dipeptide identification as their trimethylsilyl (Me_3Si) derivatives have recently been developed in previous studies of DAP sequencing techniques (Krutzsch & Pisano, 1978). Although the DAP methodology adds speed and sensitivity to N-terminal sequence analysis, it does duplicate the sense of the Edman method. A DCP method, on the other hand, would complement existing techniques, because it can be used with N-terminal blocked polypeptides and because it can facilitate the completion of sequence analyses of long polypeptides when N-terminal sequence determinations have fallen short of reaching the C terminus (Rao et al., 1979).

The present study describes a method for polypeptide sequence analysis starting from the polypeptide C terminus. This DCP C-terminal polypeptide sequencing method is based on the use of angiotensin I converting enzyme (ACE), an exopeptidase that successively removes dipeptides from the C terminus of polypeptides. A number of polypeptides, including some that were blocked at the N terminus, were digested with this enzyme, and the released dipeptides were identified as their Me_3Si derivatives by gas chromatography–mass spectrometry (GC–MS). Digestion of the polypeptides following addition or subtraction of a single C-terminal residue provided the overlapping dipeptides necessary for dipeptide alignment to give the polypeptide sequences.

Materials and Methods

Angiotensin I Converting Enzyme (EC 3.4.15.1). Angiotensin I converting enzyme (ACE), a dipeptidyl carboxypeptidase, was used to carry out polypeptide digestions starting from the C terminus. It can be purchased from Calbiochem or New England Enzyme or prepared from rabbit lung acetone powder (Sigma Chemical Co.) by recently described methods (Holmquist et al., 1979). The enzyme used here was prepared from rabbit lung acetone powder and was stored at 4 °C as a concentrated solution in a 0.05 M *N*-methylmorpholine acetate (NMM·HOAc) buffer, pH 7.2, which also contained 1% w/v NaCl. The ACE provided by Calbiochem was found to give satisfactory results, although larger amounts of this enzyme preparation were required to give activity levels comparable to the material prepared from rabbit lung.

Activity Measurement of Angiotensin I Converting Enzyme. The procedure for measurement of enzyme activity units and unit definition (Holmquist et al., 1979) will be outlined here. The substrate used in this assay was furanacryloyl-L-phenylalanylglycylglycine (FA-F-G-G). The activity of ACE on this substrate was measured by observing the loss of absorbance at 328 nm as a function of time. The assay was carried out at 37 °C in 1.0 mL of 0.05 M Tris-HCl buffer, pH 7.5, containing 0.3 M NaCl and 1×10^{-4} M FA-F-G-G. Immediately after addition of ~ 10 μL of a suitable dilution of enzyme, the absorbance reading was continuously taken, for 3–5 min. Complete digestion of this level of substrate would result in an absorbance loss of 0.230. One unit of angiotensin I converting enzyme activity is defined as that amount which digests 1 μmol of substrate/min (J. F. Rioridan, personal communication).

Digestion Conditions for Polypeptides. Polypeptides were degraded in a 0.3-mL Reacti-Vial (Pierce Chemical Co.), previously cleaned with a 1:4 v/v solution of HNO_3–H_2SO_4. Between 5 and 25 nmol of polypeptide was usually analyzed. Digestion was carried out in 50 μL of 0.05 M NMM·HOAc buffer, pH 7.2, containing 1% w/v NaCl and 0.005 unit of enzyme. The vial was sealed with a Teflon-lined cap and digestion allowed to proceed for 4 h at 37 °C before freeze-drying.

Preparation of Trimethylsilylated (Me_3Si) Dipeptides. To the vial, containing the freeze-dried digestion mixture, was added 25 μL of dry acetonitrile (Burdick and Jackson Co.) and 25 μL of *N,O*-bis(trimethylsilyl)trifluoroacetamide (BSTFA) (Supelco Co.). The vial was again tightly sealed with a Teflon-lined cap and heated at 140 °C for 10 min with occasional mixing. A 10-μL aliquot of this solution was then analyzed for Me_3Si dipeptides by gas chromatography–mass spectrometry (GC–MS).

Identification of Me_3Si Dipeptides by GC–MS. The Me_3Si dipeptides resulting from DCP polypeptide digestion were identified on an LKB Model 2091 gas chromatograph–mass spectrometer. The gas chromatograph column was a 0.6 m × 2 mm i.d. rigorously silanized glass column packed with 1% OV-1 (Pierce Chemical Co.) on 80–100 mesh Chromosorb W (Pierce Chemical Co.). Sample elution was achieved by increasing the temperature at 10 °C/min from 100 to 270 °C. Mass spectra were taken at 70 eV with a 50-mA ionizing current. The details of Me_3Si dipeptide identification have been previously described (Krutzsch & Pisano, 1978).

Modification of Polypeptides by Addition Procedure. To a 0.3-mL Teflon-capped Reacti-Vial containing the dried polypeptide sample was added 100 μL of a solution of 88% formic acid and acetic anhydride (1:1 v/v), which, after preparation, had been previously allowed to stand for 10 min at room temperature. After 2 h at room temperature with occasional mixing, the reaction mixture was evaporated to dryness with nitrogen, 50 μL of water was added, and the resulting solution was freeze-dried. To the dried formylated polypeptide sample was then added 25 μL of dry *N,N*-dimethylacetamide (DMAC) containing 0.4 mg of L-α-aminobutyric acid *tert*-butyl ester (Abu-O-*t*-Bu) as the free base and 25 μL of DMAC containing 0.2 mg of *N*-(ethoxycarbonyl)-2-ethoxy-1,2-dihydroquinoline (EEDQ) (Gold Label, Aldrich Chemical Co.). The resulting mixture was heated at 90 °C for 2 h with occasional mixing and then evaporated to dryness with nitrogen. The residue was washed twice with 100 μL of benzene and dried. The modified peptide was then deblocked at the C terminus by treatment with anhydrous trifluoroacetic acid at room temperature for 2 h (Bodansky et al., 1976), and the solution was evaporated to dryness with nitrogen. Deblocking of the polypeptide methyl and *n*-propyl esters produced in some experiments was accomplished by

heating them with 5 μL of triethylamine in 50 μL of 1:1 pyridine–water for 3 h at 90 °C. The formyl group was then removed from the modified polypeptide by treatment with 100 μL of 12 N hydrochloric acid for 2 h at room temperature with occasional mixing. The solution was evaporated to dryness with nitrogen and the residue twice freeze-dried from 50 μL of water. Prior to DCP digestion, 50 μL of 1:1 pyridine–water and 5 μL of triethylamine were added, and, after heating at 90 °C for 2 h with occasional mixing, the solution was evaporated to ~10 μL with nitrogen. Following this, 50 μL of distilled water was added, and the solution was again evaporated to ~10 μL. Finally, 50 μL of distilled water and 5 μL of pyridine were added, and the resulting solution was freeze-dried. The *tert*-butyl ester of Abu was prepared by reaction of Abu with a solution of 70% perchloric acid in *tert*-butyl acetate (Taschner et al., 1961). This compound was stored at −20 °C as its free base in a Teflon-capped vial. The methyl and *n*-propyl esters of Abu used in some experiments were prepared from reaction of Abu with thionyl chloride and methyl or *n*-propyl alcohol (Bodansky et al., 1976).

Modification of Polypeptides by Subtraction Procedure. The chemical procedure was an adaptation of the one outlined by Stark (1972). Prior to reaction, the polypeptide was formylated as described above. To the 0.3-mL Teflon-capped Reacti-Vial containing the dried formylated polypeptide sample was then added 0.4 mg of recrystallized (ethanol) ammonium thiocyanate, 50 μL of redistilled acetic acid, and 50 μL of redistilled acetic anhydride. After being heated for 2 h at 90 °C with occasional mixing, the resulting solution was evaporated with nitrogen, and the residue was washed twice with 100 μL of benzene. The formylated peptide thiohydantoin intermediate was then cleaved with 100 μL of 12 N hydrochloric acid at room temperature for 2 h with occasional mixing. This procedure removed both the C-terminal residue and the formyl group in the same step. The solution was then evaporated to dryness with nitrogen and the residue twice freeze-dried from 50 μL of water and then extracted twice with 100 μL of benzene. Prior to digestion, the shortened peptide was resolubilized in the same manner as was done in the addition procedure.

For peptides containing a C-terminal basic residue, carboxypeptidase B (CPB) was used for the subtractive procedure. The digestion of these peptides employed a modification of the method outlined by Ambler (1972). To a 0.3-mL Teflon-capped Reacti-Vial containing the dried polypeptide sample was added 50 μL of a 0.1 M NMM·HOAc, pH 7.8, buffer containing 2.0 units of CPB (Worthington Enzyme or Boehringer-Mannheim). The resulting solution was incubated for 4 h at 37 °C, then heated at 90 °C for 15 min, and finally freeze-dried.

Homoserine Lactone Containing Polypeptides. The homoserine lactone ring was opened by heating the polypeptide in 50 μL of a solution of 1.0 M NMM·HOAc, pH 7.0, in a 0.3-mL Teflon-capped Reacti-Vial at 90 °C for 1 h. The solution was then evaporated with nitrogen, 50 μL of distilled water was added, and the solution was freeze-dried. Substitution of *N*-methylmorpholine for the pyridine normally used for this conversion eliminated certain nondipeptide background peaks that otherwise would appear in the GC–MS analysis.

Results

Digestion Procedures. The conditions for obtaining maximum dipeptide yields from DCP polypeptide digestion were determined, and the enzyme preparation was tested for the presence of dipeptidase activity and for substances yielding spurious peaks upon analysis. The experimental results were determined by GC–MS analysis of the freeze-dried and trimethylsilylated (Me_3Si) digestion mixtures.

Table I: Summary of Digestion Conditions Tested

digestion condition	amounts	optimum condition
enzyme unit	0.00005, 0.0001, 0.0002, 0.0005, 0.001, 0.002, 0.005, 0.01, 0.02, 0.05	0.005 unit
buffer concn (M)	0.005, 0.01, 0.02, 0.05, 0.10, 0.20	0.05–0.10
buffer pH	6.5, 7.0, 7.5, 8.0, 8.5	7.0–7.5
digestion time (h)	0.5, 1, 2, 4, 6	4

Optimum buffer conditions for digestion were first to be determined by using Ac-Ala-Ala-Ala-Ala (1 in Table II) as substrate. The digestion was carried out in NMM·HOAc buffer at concentrations and pH shown in Table I. The optimum conditions in each case determined are also shown there. Because NMM·HOAc was found to be compatible with this system and because of its favorable p*K* and volatility, no other buffers were tested, and the NaCl concentration, required for ACE activity (Holmquist et al., 1979), was set at 1% w/v to keep the solid level low during the subsequent Me_3Si derivatization step while yielding near maximum enzyme activity.

The range of enzyme activity units and digestion time that gave optimum results were measured by using Ac-Ala-Ala-Ala-Ala as the initial test substrate. Units of ACE activity and digestion times tested are shown in Table I. Experiments performed with several polypeptides (among them, polypeptides 3 and 10 in Table II) showed that 0.0005 unit of enzyme activity, which yielded complete Ac-Ala-Ala-Ala-Ala digestion, gave incomplete digestion of these polypeptides and that 0.005 unit of activity was required. When more than 0.005 unit of enzymatic activity was used, the action of contaminating dipeptidase activity was observed. Although 2 h of digestion completely degraded Ac-Ala-Ala-Ala-Ala, test digestions with these other polypeptides showed that 4 h was the optimal digestion time. Longer times (20 h) gave lower yields of dipeptides due to their slow degradation to amino acids by the residual dipeptidase activity.

Several digestion experiments were performed in the absence of substrate to determine if spurious GC peaks arose from the digestion mixture. The gas chromatograms obtained by analysis of these enzyme and buffer mixtures exhibited no significant peaks above the base line.

Scope and Limitations of DCP Polypeptide Digestion. For establishment of the scope and limitations of the DCP method when ACE was used, a variety of polypeptides were subjected to degradation by this enzyme and the Me_3Si-derivatized digests were analyzed for the presence and yield of released dipeptides by using GC–MS. The polypeptides used in this study are shown in Table II along with the dipeptides obtained in each case. There were slight indications from these experiments that some tripeptides may be more slowly digested than longer polypeptides.

Studies on the amount of polypeptide that could be successfully analyzed by the DCP method first utilized a 49-residue polypeptide (22 in Table II), at various amounts in the range of 5–25 nmol; at all polypeptide levels, satisfactory results were obtained. The GC–MS results of the 5-nmol experiments suggested that analysis would be possible with lower amounts of material. Several other polypeptides (2, 3, and 10 in Table II) were also analyzed at the 5-nmol level and gave similarly satisfactory results. The remaining polypeptides

Table II: Polypeptides Tested with DCP Method

	polypeptide	modification	digestion products
1	Ac-Ala-Ala-Ala-Ala	none	Ac-Ala-Ala, Ala-Ala
		+Abu[a]	Ac-Ala, Ala-Ala, Ala-Abu
2	<Glu-Glu-Gln-Leu-Glu-Glu	none	<Glu-Glu, Gln-Leu, Glu-Glu[b]
3	Tyr-Gly-Gly-Phe-Met	none	Tyr, Gly-Gly, Phe-Met[b]
		+Abu	Tyr-Gly,[b] Gly-Phe, Met-Abu[b]
		−Met[c]	Tyr-Gly,[b] Gly-Phe
4	Tyr-Gly-Gly-Phe-Leu	none	Tyr, Gly-Gly, Phe-Leu
		+Abu	Tyr-Gly,[b] Gly-Phe, Leu-Abu
		−Leu	Tyr-Gly,[b] Gly-Phe
5	Phe-Leu-Glu-Glu-Leu	none	Phe, Leu-Glu, Glu-Leu[b]
		+Abu	Phe-Leu, Leu-Abu[d]
6	Pro-Phe-Gly-Lys	none	Pro-Phe, Gly-Lys
		CPB[e]	Pro, Phe-Gly
7	Glu-Gly-Phe	none	Glu, Gly-Phe
		+Abu	Phe-Abu[d]
8	Lys-Glu-Thr-Tyr-Ser-Lys	none	Lys-Glu,[b] Thr-Tyr, Ser-Lys
		CPB	Lys, Glu-Thr,[b] Tyr-Ser
9	Arg-Pro-Pro-Gly-Phe-Ser-Pro-Phe-Arg	none	Ser-Pro[f]
10	Asp-Arg-Val-Tyr-Ile-His-Pro-Phe-His-Leu	none	His-Leu
		+Abu	Tyr-Ile,[g] His-Pro,[g] Phe-His,[g] Leu-Abu
		−Leu	Tyr-Ile,[g] His-Pro,[g] Phe-His[g]
11	Pro-Ala-Pro-Glu	none	none
12	Asp-Ser-Asp-Pro-Arg	none	none
		CPB	Asp-Ser, Asp-Pro
13	Leu-Trp-Met-Arg-Phe-Ala	none	Leu-Trp, Phe-Ala[f]
14	Pro-Leu-Gly-Gly	none	Pro-Leu, Gly-Gly
15	Val-Gly-Ser-Glu	none	Val-Gly, Ser-Glu
		+Abu	Val, Gly-Ser[d]
16	<Glu-Gly-Lys-Arg-Pro-Trp-Ile-Leu	none	Ile-Leu
17	<Glu...Pro-Tyr-Ile-Leu (1 10)	none	Ile-Leu
18	Thr-Phe-Pro-Ser-Val-Arg	none	none[f]
		CPB	Thr, Phe-Pro,[b] Ser-Val
19	Phe...Thr-Pro-Lys-Ala (1 27)	none	none
20	Lys-Glu-Thr-Ala-Ala-Ala-Lys-Phe-Glu-Arg-Gln-His-Met-Asp-Ser-Ser-Thr-Ser-Ala-Ala	none	Lys-Glu,[b] Thr-Ala, Ala-Ala, Lys-Phe,[b] Gln-His, Met-Asp, Ser-Ser, Thr-Ser, Ala-Ala[f]
21	Trp-Ala-Gly-Gly-Asp-Ala-Ser-Gly-Glu	none	Trp, Ala-Gly, Gly-Asp, Ala-Ser, Gly-Glu
		+Abu	Trp-Ala,[b] Gly-Gly, Ser-Gly[d]
		−Glu	Trp-Ala,[b] Gly-Gly, Asp-Ala,[b] Ser-Gly
22	Ser...Pro-Ser-Leu-Lys-Asp-Lys-Phe-Ile-Ile-Ser-Arg-Asp-Asn-Ala-Lys-Asn-Ser-Leu-Tyr-Leu-Gln-Hse (1 28)	none	Leu-Lys,[b] Asp-Lys,[b] Phe-Ile, Ile-Ser, Asn-Ala, Lys-Asn,[b] Ser-Leu, Tyr-Leu, Gln-Hse[f]
23	Ser...Pro-Ser-Leu-Lys-Asp-Lys-Phe-Ile-Ile-Ser-Arg-Asp-Asn-Ala-Lys-Asn-Thr-Leu-Tyr-Leu-Gln-Hse (1 28)	none	Leu-Lys,[b] Asp-Lys,[b] Phe-Ile, Ile-Ser, Asn-Ala, Lys-Asn,[b] Thr-Leu, Tyr-Leu, Gln-Hse[f]
24	Phe-Ser-Trp-Gly-Ala-Glu-Gly-Gln-Arg	none	Phe, Ser-Trp, Gly-Ala, Glu-Gly[b,f]
		CPB	Phe-Ser,[b] Trp-Gly,[b] Ala-Glu, Gly-Gln
25	<Glu-Gly-Val-Asn-Asp-Asn-Glu-Glu-Gly-Phe-Phe-Ser-Ala-Arg	none	<Glu-Gly, Val-Asn,[b] Asp-Asn,[b] Glu-Glu,[b] Gly-Phe, Phe-Ser[f]

[a] Addition of L-α-aminobutyric acid to C terminus –COOH (and β- and γ-COOH). [b] Variable dipeptide loss due to dipeptidase activity. [c] Single C-terminal residue subtraction via NH_4SCN–Ac_2O. [d] The presence of β- or γ-Abu substituted dipeptides, due to presence of Asp or Glu, respectively, is not observed in GC trace. [e] Single C-terminal residue subtraction via carboxypeptidase B. [f] Arg-containing dipeptides are not observed in GC trace. [g] Detected in significantly lower amounts (on the order of approximately 4–10 times) than expected.

listed in Table II were digested in amounts varying from 25 to 200 nmol.

Digestion of the polypeptides listed in Table II showed that the amino acid side chains had no noticeable effect on dipeptide yield, whether it was on the N or C terminus of the dipeptide to be released, except when proline was encountered. As Table II shows (peptides 10, 11, 12, 16, 17, 18, 22, 23, and 24), when cleavage of a peptide bond involving the imino group of proline was required, no release of a Pro-X dipeptide occurred, and digestion terminated at that point. ACE cleavage of the peptide bond involving the carboxyl group of proline (peptides 9, 10, 12, 18, and 19 in Table II) was found to vary in efficiency among individual polypeptides.

The results from some of these digestions showed that selected dipeptides were being partially degraded by the presence of low levels of dipeptidase enzyme. Dipeptides containing N-terminal amino acid residues with charged or large side chains such as His, Lys, and Glu or Met, Phe, and Tyr, respectively, appeared to be more susceptible than others. This was evident both from a decrease in their GC peak heights relative to the other dipeptides and from the concurrent and proportional appearance in the GC trace of the Me_3Si amino acids that constituted these dipeptides. The data from analysis of a number of digestions indicated that dipeptide degradation ranged from no loss to losses of >50% in a few cases. This dipeptidase activity on a given dipeptide produced from ACE digestion of a particular polypeptide varied somewhat from digestion to digestion, possibly because of slight changes in traces of the required metal ion. However, attempts to selectively inhibit dipeptidase activity by using the metal chelating agent *o*-phenanthroline, which can be removed prior to dipeptide derivatization, were not successful, as both ACE and dipeptidase enzymatic activities appeared equally sensitive to this reagent.

As an illustration of the results obtained from DCP (ACE) digestion of a polypeptide, the gas chromatogram of the

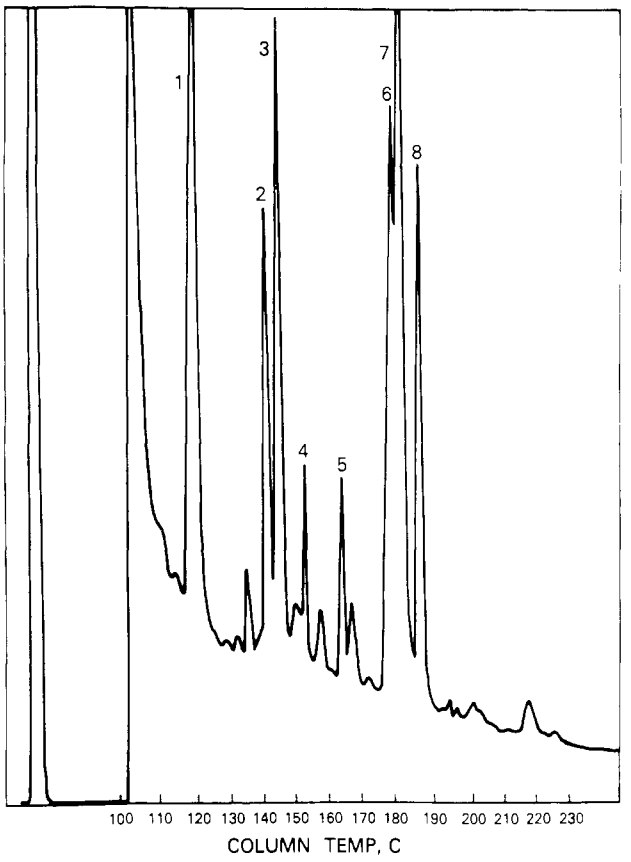

FIGURE 1: Gas chromatograph resulting from injection of 20% of the Me_3Si derivatized DCP (ACE) digestion of 25 nmol of Trp-Ala-Gly-Gly-Asp-Ala-Ser-Gly-Glu. (Peak 1) Ala-Gly and polypeptide impurity; (peaks 2, 4, and 5) polypeptide impurity; (peak 3) Ala-Ser; (peak 6) Gly-Asp; (peak 7) Trp and polypeptide impurity; (peak 8) Gly-Glu. Most of the other small unnumbered peaks are also polypeptide sample impurities.

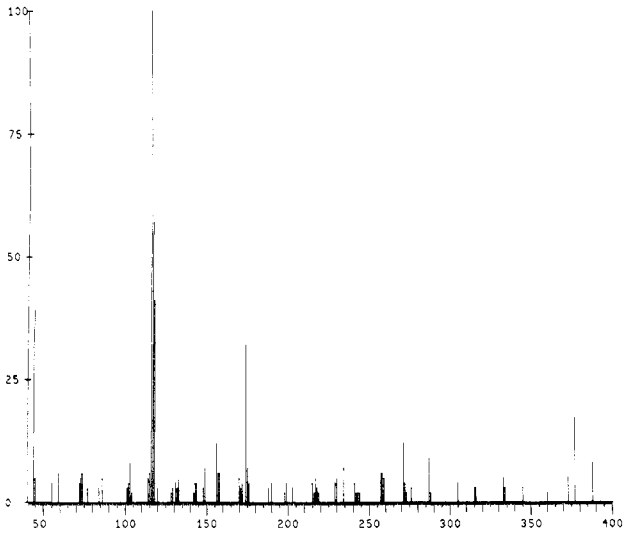

FIGURE 2: Mass spectrum of Ala-Ser (peak 3 in Figure 1). Peak at m/e 116 is the sequence-determining ion for N-terminal Ala; peak at m/e 377 is M − 15 ion, yielding dipeptide molecular weight for Me_3Si-Ala-Ser.

Me_3Si-derivatized digestion mixture of 25 nmol of polypeptide 21 is shown in Figure 1. The extraneous peaks were demonstrated to be polypeptide contaminants. Their mass spectra indicated that they were of nonpeptide nature; no further attempts were made at their characterization. The mass spectrum of one of the dipeptides, Ala-Ser, is shown in Figure 2.

Table III: Summary of Conditions Tried for Making Modified Polypeptides

reaction condition	variations investigated	best variation
	(A) Single Add-On	
coupling agents (CA)	EEDQ, CDI	EEDQ
Abu derivative (AD)	Abu-O-Me, Abu-O-Pr, Abu-O-*t*-Bu	Abu-O-*t*-Bu
reagent ratios (CA/AD)	1:2, 1:1, 2:1	1:2
temp (°C)	25, 50, 70, 90, 110	90
	(B) Single Subtraction	
acid/acid anhydride(s)	acetic/acetic, acetic/acetic, trifluoroacetic, trifluoroacetic/trifluoroacetic	acetic/acetic
temp (°C)	50, 70, 90, 110	90

Preparation of Modified Polypeptide. Methods were investigated to generate the modified polypeptide by either adding or subtracting a single residue at the C terminus of the original polypeptide. When DCP digested, the modified polypeptide provided the dipeptides that allow alignment of the dipeptides released from DCP digestion of the native polypeptide. The effect of reaction conditions was determined by DCP digestion of the modified polypeptide and GC–MS analysis of the digestion products. The reaction yield was determined by observing if any of the dipeptides from the unmodified peptide were produced from the DCP digestion. The yield of the modified polypeptide available for DCP digestion was determined by observing how much of the dipeptides from the modified polypeptide was obtained from the digestion.

The addition method for generation of the modified polypeptide, which involved coupling of one residue of L-α-aminobutyric acid to the polypeptide C terminus, was studied first with Ac-Ala-Ala-Ala-Ala. The conditions and reagents tested for this procedure are shown in Table III, as are the optima determined in each case. The add-on reaction was then tested with several polypeptides possessing differing C-terminal residues, and DCP digestion of the modified polypeptides (1, 3, 4, 5, 7, 10, 15, and 21 in Table II) showed that the coupling reaction converted 95–100% of the polypeptide into the modified polypeptide. The yields of dipeptides from DCP digestion of these modified polypeptides were approximately 80–90%. Released dipeptides obtained from DCP digestion of several polypeptides containing Asp or Glu (5, 7, 15, and 21 in Table II), following addition of Abu, showed that the β- or γ-Abu derivatives of these residues did not block the action of ACE. Several experiments conducted with polypeptides containing a C-terminal homoserine lactone group showed that peptides containing this residue at the C terminus would also require a coupling agent for addition of Abu.

The modified polypeptide can also be prepared by removing one amino acid residue from the C terminus of the original polypeptide. Several procedural changes were required to adapt the published procedure (Stark, 1972) to the DCP system, using methionine enkephalin as a test polypeptide (3 in Table II). The reaction conditions studied are shown in Table III, as are the optima determined in each case. Several other polypeptides (4, 10, and 21 in Table I) were then tested, and the results of those experiments showed that 95–100% of the dipeptides resulting from DCP digestion were found to be from the modified polypeptide, and yields of released dipeptides were approximately 80–90%. When the modified polypeptide was obtained by this means, the β- and γ-carboxyl groups of

Asp and Glu were unaltered.

For sequence determination of peptides with Arg or Lys at the C terminus, carboxypeptidase B (CPB) can be used to generate the modified polypeptide via a single subtraction step. This approach was studied to determine how many units of CPB activity were needed and if any interfering GC peaks were produced. The polypeptide Lys-Glu-Thr-Ser-Lys (8 in Table II) was used as the test substrate. These experiments showed that ~2 units of CPB was optimal and that no spurious GC peaks came from this enzyme treatment. Again, 95–100% of the dipeptides observed after DCP digestion came from the modified polypeptide and were produced from the digestion in ~90% yields based on the amount of polypeptide originally present. Similar results were also obtained with several other polypeptides (6, 12, 18, and 24 in Table II) with C-terminal Lys or Arg when they were sequentially treated with CPB and DCP.

Discussion

This report outlines conditions developed for an enzymatic procedure for polypeptide sequence analysis starting from the C terminus using an approach based on digestion with dipeptidyl carboxypeptidase. With a few exceptions, the method appears to be generally applicable, and small amounts of polypeptide can be analyzed. The DCP method complements the widely used N-terminal Edman degradation as well as the dipeptidyl aminopeptidase methods for polypeptide sequence analysis.

The DCP procedure obtains polypeptide sequences by degradation of the polypeptide into dipeptide fragments. These dipeptides are then aligned into the correct sequence with the information provided by the dipeptides obtained from DCP digestion of the polypeptide after a single C-terminal residue has been either added or removed:

original polypeptide

$$\text{A-B-C-D-E-F} \xrightarrow{\text{DCP}} \text{A-B + C-D + E-F}$$

modified polypeptide

$$\text{A-B-C-D-E-F-G} \xrightarrow{\text{DCP}} \text{A + B-C + D-E + F-G}$$

(single addition)

$$\text{A-B-C-D-E} \xrightarrow{\text{DCP}} \text{A + B-C + D-E}$$

(single subtraction)

A possible alternative to the use of addition or subtraction procedures for alignment of the dipeptides from DCP digestion of the polypeptide might come from homology comparisons or from a time release study.

The first step in the overall DCP procedure involves enzymatic production of dipeptides. Efficient digestion requires that the dipeptides be produced in high yield to give maximum sensitivity, that the enzyme preparations contain no extraneous proteolytic activity which would either give confusing results or further degrade the enzyme and that no material be introduced which would give rise to spurious peaks in the GC–MS analysis. This work was simplified by the work of others (Holmquist et al., 1979; J. F. Rioridan, personal communication), who established overall pH and NaCl requirements for ACE and showed that the angiotensin I converting enzyme obtained by their procedure was free of endopeptidase, aminopeptidase, and carboxymonopeptidase activity. However, successful application of the DCP methodology required studies of ACE with respect to compatibility with volatile buffer, optimization of buffer pH and concentration, enzyme units, and digestion times. It was also necessary to determine the presence of dipeptidase activity and of materials yielding spurious GC peaks during dipeptide analysis. The successful digestion of a variety of polypeptides into dipeptides, as discussed below, indicated that the level of enzymatic units and digestion times chosen adequately overcame the probable variation in ease of dipeptide cleavage from the polypeptide and possible product inhibition by released dipeptides.

The digestive action of ACE on a variety of polypeptides to degrade them into dipeptides was then investigated by using the optimum digestion conditions defined by these studies. As Table II shows, polypeptide chain length presented no barrier to digestion by ACE, since polypeptides varying from 3 to 49 residues were successfully attacked by this DCP enzyme. While the 49-residue polypeptide was not fully digested (due to blockade by Pro), the datum still shows that a polypeptide of this size is accessible to DCP (ACE) digestion. The data in Table II also show that the enzyme was found to have a broad range of digestive activity on polypeptides, releasing dipeptides containing acidic, basic, hydrophilic, and hydrophobic side chains. Although it was reported (Elisseeva et al., 1971) that ACE is inert toward substrates with C-terminal Glu, the results with polypeptides containing this residue at the C terminus (polypeptides 2, 15, and 21 in Table II) showed that ACE also had activity on this class of polypeptides. Peptide bonds involving the imino group of proline were not cleaved by this enzyme, and digestion stopped at that point. This fact has been previously noted in other work (Bakhle, 1974). Peptide bonds involving the carboxyl group of proline are cleaved, but not in all cases, and cleavage at the C terminus of Pro to release X-Y from a Pro-X-Y sequence appears to depend on the polypeptide chain attached to the N terminus of the proline moiety. For example, the yield of dipeptides from the oxidized insulin B chain (polypeptide 19 in Table II), was very low, in the case of bradykinin (polypeptide 9 in Table II), the second carboxyl proline cleavage did not occur, and with the modified polypeptide derivative of angiotensin I (with either single add-on or subtraction) (polypeptide 10 in Table II), only low amounts of cleavage of His-Phe from Pro-His-Phe were observed. It is conceivable that a particular polypeptide without Pro could have a conformation that would hinder access to the C terminus by ACE, which would again cause low dipeptide yields. There is also the possibility that the intact polypeptide, or a partially degraded oligomer, would not bind to ACE and thus would not be a substrate at one of these possible points.

The angiotensin-converting enzyme preparation used here was found to contain a low level of dipeptidase activity, causing varying losses of certain dipeptides as previously described. This could decrease somewhat the ability to detect those dipeptides on which it was more active, such as those containing N-terminal amino acids with charged or larger side chains such as Lys, His, and Glu or Met, Phe, and Tyr, respectively. However, the method was found to be sensitive; by use of the 49-residue polypeptide (22 in Table II) as a test substrate, levels down to 5 nmol gave easily detectable and identifiable levels of dipeptides. A strong suggestion that the 1-nmol level could be reached was exhibited by these gas chromatograms. The scope and sensitivity of the DCP method discussed here is comparable to the previously described DAP method (Krutzsch & Pisano, 1977; Seifert & Caprioli, 1978).

The shortcomings of the current DCP enzyme preparations could be circumvented in several ways. For example, the low level of dipeptidase activity in the ACE preparation can probably be decreased further or eliminated by further pu-

rification. The problem of imino proline cleavage could be eliminated by discovery of another DCP enzyme which may specifically cleave that particular peptide bond. The possibility that some susceptible bonds, especially those involving the carboxyl group of Pro, may not be cleaved by ACE, could be eliminated by using other DCP enzymes. A DCP enzyme isolated from *E. coli* (Yaron, 1976), shown to cleave Gly-Phe from Arg-Pro-Pro-Gly-Phe (which ACE cannot do), represents such a possibility.

Several methods were investigated to generate the modified polypeptides and involved testing strategies for both addition or subtraction of one residue at the polypeptide carboxyl terminus. The addition of Abu to the polypeptide C terminus was an effective step for obtaining the modified polypeptide, and the C-terminal residue was also labeled. However, as expected, the side chains of Asp and Glu also reacted and, following digestion, dipeptides containing Asp and Glu appeared as the corresponding β- or γ-Abu-substituted dipeptides, which, like Arg-containing dipeptides, would have to be identified by direct probe introduction, rather than by gas chromatographic introduction into the GC–MS. This manipulation is required because irreversible adsorption to the GC column packing occurs. However, because of their more highly polar nature, these Me_3Si derivatives begin to volatilize on the probe after most of the other Me_3Si dipeptides, allowing a separation to be effected. The sensitivity of this detection is similar to that of the usual GC–MS analysis. Unless a large number of these residues were present in the polypeptide, this requirement would not be a problem. The action of ACE was not inhibited by the presence of these Abu-substituted side-chain groups in the polypeptide.

Two subtractive methods were explored for generating the modified polypeptide. One of these was a chemical method, using an alteration of the one detailed by Stark (1972). This adapted procedure satisfactorily subtracted a single C-terminal residue from each of the polypeptides on which it was tested (3, 4, 10, and 21 in Table II). In addition, this reaction did not alter Asp or Glu, allowing analysis of dipeptides containing these residues to be carried out in the usual way. The other subtractive method, which could be used only with polypeptides that had a basic C-terminal residue, involved digestion with carboxypeptidase B. Little or no GC background-forming materials arose from this enzyme; in particular, extraneous dipeptides resulting from its possible digestion by DCP were not observed. Because all activity in the CPB preparation was denatured prior to the addition of DCP for the second digestion step, problems due to contaminating dipeptidase or monopeptidase activity in this preparation or to removal of an internal basic residue exposed during the course of subsequent DCP digestion were eliminated. At this point, no conclusion has been reached concerning which chemical (add-on or subtraction) method has the greatest general utility for generating the modified polypeptide. In the case of polypeptides with a basic C-terminal residue, it is evident that CPB digestion was clearly the most facile means for obtaining the modified polypeptide.

The versatility of the DCP technique is further enhanced by the number of alternative means available for dipeptide identification and should be useful in many laboratories. Other means for characterizing the dipeptide products from DCP or DAP polypeptide digestion might include ion-exchange chromatography (Callahan et al., 1970) and thin-layer chromatography (Paukovits, 1973). In our studies, the GC–MS method has proven to be a rapid, simple, and sensitive means for dipeptide analysis. Both the DAP and DCP methods have been facilitated by computer programs recently written that contain automated procedures for dipeptide identification and for dipeptide alignment.

In summary, the dipeptidyl carboxypeptidase system for polypeptide sequence analysis can provide a valuable addition to the tools now available for polypeptide sequence analysis. The DCP procedure complements both the Edman method and the dipeptidyl aminopeptidase method for sequence analysis, and, like the DAP method, it has advantages of speed, simplicity, and sensitivity. For sequence analysis of polypeptides that have a blocked N terminus, DCP digestion should become the method of choice.

Acknowledgments

The author acknowledges the generous help provided by Dr. Thomas J. Kindt in preparing this manuscript.

References

Ambler, R. P. (1972) *Methods Enzymol. 25*, 262.
Bakhle, Y. S. (1974) *Handb. Exp. Pharmakol. 37*, 41.
Bodansky, M., Klausner, Y. S., & Ondetti, M. A. (1976) *Peptide Synthesis*, 54.
Callahan, P. X., Shepard, J. A., Reilly, T. J., McDonald, J. K., & Ellis, S. (1970) *Anal. Biochem. 38*, 330.
Edman, P. (1970) *Mol. Biol., Biochem. Biophys. 8*, 211.
Elisseeva, Y. E., Orekhovich, V. N., Pavlikhina, L. V., & Alexeenko, L. P. (1971) *Clin. Chim. Acta 31*, 413.
Holmquist, B., Bunning, P., & Riordan, J. F. (1979) *Anal. Biochem. 95*, 540.
Krutzsch, H. C., & Pisano, J. J. (1977) *Methods Enzymol. 47*, 391.
Krutzsch, H. C., & Pisano, J. J. (1978) *Biochemistry 17*, 2791.
Paukovits, W. R. (1973) *J. Chromatogr. 85*, 154.
Rao, D. N., Rudikoff, S., Krutzsch, H., & Potter, M. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 2890.
Seifert, W. E., Jr., & Caprioli, R. M. (1978) *Biochemistry 17*, 436.
Stark, G. R. (1972) *Methods Enzymol. 25*, 369.
Taschner, E., Chimiak, A., Bator, B., & Sokolowska, T. (1961) *Justus Liebigs Ann. Chem. 646*, 134.
Yaron, A. (1976) *Methods Enzymol. 45*, 599.