THE AMINO ACID SEQUENCE OF PORCINE THYROCALCITONIN

BY J. T. POTTS, JR., H. D. NIALL, H. T. KEUTMANN, H. B. BREWER, JR., AND L. J. DEFTOS

NATIONAL HEART INSTITUTE, NATIONAL INSTITUTES OF HEALTH, BETHESDA, MARYLAND

Communicated by C. B. Anfinsen, February 21, 1968

Thyrocalcitonin, the recently discovered thyroid hormone that causes hypocalcemia,^{1, 2} has been prepared in homogeneous form and shown to be a singlechain polypeptide composed of 32 amino acids.³⁻⁵ The amino acid composition and specific biological activity of thyrocalcitonin isolated from acid extracts of porcine thyroid tissue are similar to that of the hormone which we had earlier purified from thyroids extracted with urea.^{3, 5, 6} Reports from several laboratories indicate complete agreement on the amino acid analysis of thyrocalcitonin whether isolated after extraction with acid or urea.^{$7-9$}

We have determined the amino acid sequence of thyrocalcitonin through the use of sequential degradation by the phenylisothiocyanate procedure, digestion with exopeptidases, and isolation and analysis of peptide subfragments resulting from cleavage of the polypeptide by trypsin, pepsin, chymotrypsin, papain, cyanogen bromide, limited hydrolysis in dilute acid, and limited hydrolysis in 5.7 N HCL. A unique solution for the amino acid sequence was obtained; complete agreement was seen in the results from the several independent approaches. This report outlines the methods used and the results achieved during this structural analysis of thyrocalcitonin.

Materials and Methods.—Procedures used in purification of thyrocalcitonin and tests for homogeneity of the polypeptide are described in detail in other reports. $3-5$ Amino acid analyses were made with automatic analyzers equipped for high sensitivity and rapid elution analysis.^{10, 11} During acid hydrolysis in 5.7 N HCl, mercaptoethanol was usually added at a concentration of $1-2000 \, (v/v)$. This technique improves recovery of certain amino acid residues such as methionine and tyrosine.¹² Total enzymic hydrolysis of thyrocalcitonin and peptide subfragments was performed using a combination of papain and aminopeptidase M. This newer technique, involving a 5-hr incubation in $0.02 \, M$ mercaptoethanol, has been shown to provide accurate information concerning the content of asparagine, glutamine, tryptophan, and methionine or methionine sulfoxide in a protein or polypeptide.13 Asparagine, glutamine, aspartic acid, and glutamic acid were distinguished on the amino acid analyzers by use of the lithium citrate buffer systems.¹⁴ Standard techniques were used for performic acid oxidation,¹⁵ reduction and alkylation,¹⁶ and titration of thiol groups.'7

Phenylisothiocyanate degradations were performed by the three-stage procedure of Edman.'8,19 Considerable modification of the extraction schedules was made when smaller peptides were degraded in order to avoid undue losses of material. These involved reduction in the extent of benzene extraction after coupling, omission of ethyl acetate extraction, and, in a few cases, substitution of sublimation at 60° in high vacuum for benzene extraction.²⁰ The results of sequential Edman degradation were evaluated by direct analysis of the phenylthiohydantoin derivatives of successive amino-terminal amino acids (PTH-amino acids). This was accomplished principally by gas-liquid chromatography²¹ and thin-layer chromatography.¹⁹ The gas chromatographic technique provided rapid detection of the PTH-amino acids at high sensitivity and permitted quantitation of each PTH-amino acid formed; yields could be calculated thereby for each step in the degradation. Multiple criteria were used for identification of certain p)henylthiohy(Iantoin amino acids; for example, the phenyithiohydantoin of S-carboxymethylcysteine (PTH-SCMC) was also identified by the mass spectrographic pattern (compared with that of ^a standard of PTH-SCMC) and by autoradiography after thinlayer analysis'9 of the phenylthiohydantoin product formed from thyrocalcitonin labeled with ^C'4-iodoacetic acid. Reductive cleavage by sodium metal in liquid ammonia was lerformed in the absence of a proton donor.22

Endopeptidase-free carboxypeptidase and leucine aminopeptidase (the latter prepared by the method of Himmelhoch and Peterson²³) were used as described previously.²⁴ Digestions with trypsin and chymotrypsin were monitored in ^a pH-stat at pH 8.0. Trypsin was treated with diisopropylfluorophosphate to eliminate chymotrypsin contamination.²⁵ Digestion with pepsin was performed for 2 hr in 0.5 N formic acid; treatment with papain was for 8 hr in $0.05 M \text{ NH}_4$ Ac, pH 5.35. Limited hydrolysis in dilute acid was performed by heating thyrocalcitonin at a concentration of 1 mg/ml in 0.03 N HCl at 110° for periods of 6,15, and 20 hr.²⁶ Limited hydrolysis in 5.7 HCl was accomplished by heating in vacuo for 30 min at 100° .²⁶

Peptide fragments resulting from the various methods of cleavage were separated by two-dimensional electrophoresis chromatography on filter paper,25 gel filtration on Sephadex G-25 (superfine), or Bio-Gel P-2, or ion-exchange chromatography on columns of Dowex-50.27 Peptides detected in the effluent from the ion-xchange columns were assessed for purity by chromatography on thin-layer plates of cellulose; the thin-layer l)lates were sprayed with ninhydrin, Erhlich's, Sakaguchi, and Pauly reagents. Radioactive peptides (from digests of thyrocalcitonin labeled with $C¹⁴$ -iodoacetic acid) were detected by autoradiography.

Results and Discussion.—Amino acid composition and chemical properties of thyrocalcitonin: Thyrocalcitonin, after an over-all purification of 50,000-fold, had been monitored to ^a stage of constant specific activity and amino acid composition during repetitive column fractionations. $3-5$ The peptide was shown to be homogeneous on thin-layer chromatography in several solvent systems and gave a single band in disc-gel electrophoresis⁶ and a single amino-terminal acid on end-group analysis $\frac{1}{2}$ -cystine detected as cysteic acid or S-carboxymethylcysteine by the Edman procedure).^{3, 5} The amino acid composition of the purified peptide was determined after acid and enzymic hydrolysis. The combined results indicated that the amino acid composition of thyrocalcitonin was Arg₂, His₁, Asn₄, Thr₂, Ser₄, Glu₁, Pro₂, Gly₃, Ala₁, Val₁, Met₁, Leu₃, Phe₃, Tyr₁, Trp₁, $\frac{1}{2}$ Cys₂ (analyzed as cysteic acid or S-carboxymethylcysteine). Isoleucine, lysine, aspartic acid, or glutamine were not detected. A number of tests were made to ensure that the isolated peptide did indeed represent thyrocalcitonin. No evidence was found for a biologically active trace component or cofactor noncovalently bound to thyrocalcitonin. The purified product was subjected to fractionation by partition in organic solvents and preparative gel electrophoresis in S-molar urea; all biological activity remained associated with the peptide.^{3, 5} The absence of amino sugars, other amino acid derivatives, or unusual peptide linkages was indicated by the results obtained after enzymic digestion. Yields of all amino acids obtained by enzymic digestion were in complete agreement with the findings after acid hydrolysis and no unusual amino acid peaks were detected on the analyzers. The state of the sulfhydryl containing amino acids in the molecule has been extensively investigated and is being reported elsewhere.4 There was no evidence of free thiol groups; no reaction with Ellman's reagent¹⁷ was seen (even after prior treatment of the hormone with 0.1 M ethylenediaminetetraacetic acid to remove any bound metals); no S-carboxymethylcysteine

was detected after treatment of the intact polypeptide with a 50-fold molar excess of iodoacetic acid. On the other hand, after reduction with mercaptoethanol, addition of iodoacetic acid resulted in formation of two moles of S-carboxymethylcysteine per mole; two moles of free thiol were detected with the Ellman reagent. To eliminate the possibility of interchain disulfide bridges, performic acid-oxidized thyrocalcitonin was subjected to gel filtration on Sephadex G-25. The Kd (0.4) observed for performic acid-treated thyrocalcitonin was similar to that of the native polypeptide (0.5). This change, if significant, indicated only a slight increase in apparent molecular volume rather than a decrease and showed that the two $\frac{1}{2}$ -cystine residues are present as an intrachain disulfide rather than as interchain bridges.

Although there have been reports that the amino-terminal amino acid has a blocked amino group,7 we found that thyrocalcitonin has a free alpha-amino group and identified the amino-terminal amino acid by several techniques. When reduced and alkylated thyrocalcitonin was examined by the Edman procedure, a high yield of PHT-SCMC was identified by multiple criteria (see Methods). Extensive amino acid release occurred after incubation of alkylated or oxidized thyrocalcitonin with endopeptidase-free leucine aminopeptidase: S-carboxymethylcysteine or cysteic acid, respectively, were rapidly released in high yield. On the other hand, no amino acids were released after incubation of thyrocalcitonin with carboxypeptidase A or B. It seemed that this might be due to blocking of the alpha-carboxyl group (i.e., an alpha-carboxyl amide) or, since proline was present in the peptide, the presence of proline at the carboxylterminus of the molecule.

Trypsin digestion: Digestion of thyrocalcitonin with trypsin, monitored in the pH-stat, revealed a rapid digestion of reduced and alkylated or performic acid-oxidized thyrocalcitonin but a slow rate of digestion of the native molecule. Base uptake was somewhat higher than theoretical for cleavage of two peptide bonds (two arginines). Examination of tryptic digests by thin-layer chromatography showed three principal peptides and one or more minor peptides. Secondary, chymotrypticlike cleavages occurred before complete cleavage at the two arginines; however, this did not prevent isolationof the three major subfragments. Three peptides (termed $T_1, T_2,$ and T_3) were obtained in pure form after fractionation on Dowex-50. Amino acid analysis after acid hydrolysis indicated the following compositions (expressed as mole fractions): T_1 : Arg 0.08, SCMC 0.11, Asp 0.11, Thr 0.08, Ser 0.24, Ala 0.08, Val 0.09, Leu 0.17, Tyr 0.05; T_2 : Arg 0.13, His 0.14, Asp 0.43, Leu 0.14, Phe 0.15; T_3 : Thr 0.09, Ser 0.08, Glu 0.09, Pro 0.18, Gly 0.29, Met 0.08, Phe 0.18 (one mole of Trp per mole of T_1 identified after enzymic hydrolysis). T_2 and T_3 were analyzed in duplicate (two separate column fractionations); T_1 was not obtained in satisfactory yield from the second peptide fractionation. In general, peptides T_2 and T_3 were uniformly obtained in high yield but recovery of T_1 was difficult. We found that the site of chymotrypticlike cleavages was within the region of the polypeptide corresponding to tryptic peptide T_1 by examining the unfractioned tryptic digest by thin-layer chromatography, Edman degradations, and digestions with exopeptidases.

The amino acid composition of the peptides suggested that the 14-amino acid

peptide containing arginine and two moles of S-carboxylcysteine/mole (T_1) represented the amino-terminal peptide of the molecule, the 7-amino acid peptide containing arginine (T_2) represented the middle tryptic peptide, and the 11amino acid peptide, devoid of a basic residue (T_3) , represented the carboxyterminal tryptic peptide.

Cleavage with cyanogen bromide: It was necessary to prepare peptide subfragments by several additional means of cleavage to provide the necessary information to align the tryptic subfragments. After cleavage with cyanogen bromide, the lyophilized digest was fractionated on Sephadex G-25. A peak with aromatic absorbance eluted early and a second peptide was detected near the salt volume of the column. Amino acid analysis of the second fraction indicated that it was heptapeptide: Gly₂, Pro₂, Phe₁, Glu₁, Thr₁. The composition of the large peptide fragment corresponding to the peak of aromatic absorbance confirmed that it consisted of tryptic peptides T_1 and T_2 plus four additional residues: Phe₁, Ser₁, Gly₁, Homoserine₁. The latter residues apparently represented the amino-terminal portion of the carboxyl-terminal tryptic peptide. This supported the impression that the order of the tryptic peptides, beginning at the amino-terminus, was T_1, T_2, T_3 . This formulation is shown in Figure 1.

Sequential Edman degradation: Sequential Edman degradation was performed with the following peptides: (1) intact thyrocalcitonin, (2) tryptic peptides T_2 and T_3 , and (3) the heptapeptide resulting from cyanogen bromide cleavage. The sequence of the entire molecule was determined by this technique (Fig. 1). The longest repetitive degradation (on the intact molecule) involved 16 cycles. Degradation of this fragment proceeded through the region of the molecule corresponding to tryptic peptide T_1 and T_2 , thereby showing that peptide T_1 was amino-terminal and T_2 the middle tryptic peptide. Degradation of the carboxylterminal peptide was continued into the region corresponding to the heptapeptide derived by cyanogen bromide cleavage (Fig. 1). The overlapping of the repetitive Edman degradations (residues 15-16 and 26-27) further strengthened the assignment of sequence positions. Average yields for successive steps during degradation of each of the peptide fragments were: (1) intact thyrocalcitonin (residues 1 through 16), 93 per cent; (2) tryptic peptide T_2 (residues 15 through 21), 86 per cent; (3) tryptic peptide T_3 (residues 22 through 27), 85 per cent; (4) cyanogen bromide heptapeptide steps 1-3 (residues 26 through 28), 70 per cent; and steps 4-7 (residues 29 through 32), 92 per cent.

The somewhat lower yields for the shorter peptides were due mainly to losses of material during extractions. This was shown by the results obtained with the cyanogen bromide heptapeptide. Noting repetitive yields as low as 70 per cent for steps ¹ through 3, we reduced the extent of extraction (see Methods); yields for steps 4-7 rose to 92 per cent. The very high efficiency of cleavage at each step in the degradation, even with the shorter peptides, was indicated by subtractive analysis. After removal of the first six residues (15 through 20) of the T_2 heptapeptide, an aliquot of the residue was analyzed on the amino acid analyzer directly and after acid hydrolysis; only arginine (residue 21) was detected in greater than 90 per cent yield. By a similar analysis, the yield of the final residue of the cyanogen bromide heptapeptide (residue 32) was found to be 96 per cent.

MET GLY PHE

 \tilde{c}^* ម្លុះ $\ddot{\tilde{\pi}}$

 $\begin{bmatrix} 4R_0 & 4S_N \\ -I_0 & 4S_N \\ 4S_N & 4S_N \end{bmatrix}$

管" \mathfrak{F}^{α} \tilde{z}]

 \tilde{g}^2 곱**=1** 곱=1

 $\mathfrak{F}^{\mathsf{b}}$

 $\tilde{\Xi}^{\circ}$ SE^{α}

Cys SER ASN LEY SI

EDMAN
DEGRADATIONS

 $\vec{\mathsf{s}}^\circ$

 $\ddot{}$

Residue 32, the carboxyl-terminal amino acid of thyrocalcitonin was found to be prolinamide rather than proline. Thus, the resistance to carboxypeptidase degradation (found with native thyrocalcintonin, tryptic peptide T_3 , and the cyanogen bromide heptapeptide) had a dual explanation; proline was COOHterminal and was present as the alpha-carboxyl amide. Direct analysis of the residual cyanogen bromide heptapeptide after six cycles of Edman degradation on the amino acid analyzer revealed no proline but a high yield of a ninhydrin positive compound eluting at 41 minutes on the basic column. Authentic prolinamide eluted at the identical position; the unknown and reference compounds had a unique ratio of absorbance at $570-440$ m μ , 1.14. Treatment of the residue of the cyanogen bromide heptapeptide with leucine aminopeptidase caused a disappearance of the prolinamide peak and produced an equivalent yield of proline, confirming the identification of residue 32 as prolinamide. Further, cleavage of intact thyrocalcitonin with sodium metal in liquid ammonia gave a high yield of prolinamide (as expected, by reductive cleavage at the amino-terminus of a carboxyl-terminal prolinamide).

Sequential digestion of intact thyrocalcitonin and tryptic peptide T_3 by leucine aminopeptidase and tryptic peptides T_1 and T_2 by carboxypeptidase indicated an order of release of amino acids completely consistent with the proposed sequence.

Peptide subfragments from enzymic and limited acid hydrolysis: The data obtained by Edman degradation provided the solution of the complete amino acid sequence of the molecule, but it was felt important to obtain additional information by completely independent methods. Multiple peptides were obtained by digestion with pepsin and chymotrypsin and by limited hydrolysis in 0.03 N HCl. These procedures, as well as cleavage by cyanogen bromide and trypsin, provided at least five unique sets of peptide fragments that correspond to the composition of the entire molecule (Fig. 1).

Noting that these specific techniques did not provide frequent cleavages for the carboxyl-terminal portion, residues 29 through 32, and the amino-terminal region, residues ¹ through 9, two other means of cleavage were employed (papain and limited hydrolysis in 5.7 N HCl), aimed at providing multiple overlapping fragments of the terminal portions of the molecule. To select fragments from only these areas, we used thyrocalcitonin treated with C'4-iodoacetic acid (to mark peptides derived from the immediate vicinity of the $\frac{1}{2}$ -cystine residues 1 and 7), and the cyanogen bromide heptapeptide.

These combined techniques provided 59 peptide subfragments of the molecule, of which 8 peptides were obtained from more than one method. The composition of each of these 59 fragments was consistent with the sequence based on sequential Edman degradation. Furthermore, the distribution of cleavages within the molecule allowed the absolute sequence to be established, save for the relative order of residues 7-8, 8-9, and 23-24 (Fig. 1). Peptide fragments equivalent to cleavages at these positions were obtained by another series of Edman degradations. Seven- and eight-step degradations were performed on two aliquots of the intact thyrocalcitonin molecule, and a two-step degradation on tryptic peptide T3. Each of the resulting shortened peptides was purified and

shown to be homogeneous by thin-layer chromatography. Analysis gave the expected compositions for peptides consisting of residues 8 through 32, 9 through 93, and 24 through 32 (dashed arrows, Fig. 1). Adding the composition of these three peptides to the composition of the 51 peptides shown in Figure ¹ completed the information necessary to give a second, independent solution of the amino acid sequence of thyrocalcitonin.

Summary.-The covalent structure and complete amino acid sequence of porcine thyrocalcitonin has been deduced. Proof that this proposed structure is correct requires synthesis of the peptide followed by demonstration that the synthetic product has full biological and immunological activity. However, we obtained a unique solution for the sequence by several independent methods. All 32 amino acids in the polypeptide were positioned through repetitive degradations by the Edman procedure. The phenylthiohydantoin derivatives were identified by multiple criteria; high yields of the cleavage products at each step were achieved as shown both by direct analysis of the PTH-amino acids formed and by subtractive analysis of the residual peptide after degradation. A second approach involved the separation and analysis of 59 peptide subfragments of the molecule produced by enzymic and limited acid hydrolysis. Analysis of these numerous peptides was in total agreement with the sequence deduced from the Edman degradations; in fact, this second approach also provided a unique solution for the amino acid sequence identical to that based on the Edman procedure. This proposed sequence was completely supported by the results found after digestion of the native polypeptide and large peptide subfragments by exopeptidases.

The thyrocalcitonin peptide has several distinctive structural features. The single tyrosine and tryptophan are adjacent; the proportion of charged amino acids is low. There is a 1-7 intrachain disulfide bridge, providing a 23-membered ring at the amino-terminus. The carboxyl-terminal amino acid is prolinamide. The terminal disulfide loop and the alpha-carboxyl amide are analogous to structural features found in oxytocin and vasopressin. The portions of this structure important for biological and immunological activity are not yet known; however, we have found that methionine, residue 25, may be alkylated or oxidized without loss of biological activity of the thyrocalcitonin derivative.⁴

The authors gratefully acknowledge Miss Rosemary Ronan, Miss Bess Dawson, and Mr. Lynn Callahan for their expert technical assistance, and Miss Susan Siegel for her lpreparation of the manuscript; Dr. John Pisano for his collaborative assistance in the application of the gas chromatographic techniques; and Dr. Peter Gordon for his assistance with the mass spectrographic analyses. We wish to express our gratitude to Dr. Robert Schlueter, Dr. James Bastian, and their associates at Armour Pharmaceutical Company, Kankakee, Illinois, for providing the partially purified thyrocalcitonin which made this structural analysis possible.

¹ Munson, P. L., and P. F. Hirsch, Am. J. Med., 43, 678 (1967).

² Copp, D. H., Am. J. Med., 43, 648 (1967).

³Potts, J. T., Jr., H. B. Brewer, Jr., R. A. Reisfeld, P. F. Hirsch, R. Schlueter, and P. L. Munson, in Parathyroid Hormone and Thyrocalcitonin (Calcitonin), ed. R. V. Talmage and L. F. Belanger (New York: Excerpta Medica, in press).

⁴ Brewer, H. B., Jr., H. T. Keutmann, R. A. Reisfeld, R. Schlueter, P. L. Munson, and J. T. Potts, Jr., Federation Proc. (1968), in press.

⁶ Brewer, H. B., Jr., R. A. Reisfeld, R. Schlueter, P. L. Munson, and J. T. Potts, Jr., submitted to J. Biol. Chem.

⁶ Potts, J. T., Jr., R. A. Reisfeld, P. F. Hirsch, A. B. Wasthed, E. F. Voelkel, and P. L. Munson, these PROCEEDINGS, 58, 328 (1967).

⁷ Putter, I., E. A. Kaczka, R. E. Harman, E. L. Rickes, A. J. Kempf, L. Chaiet, J. W. Rothrock, A. W. Wase, and F. J. Wolf, J. Am. Chem. Soc., 89, 5301 (1967).

⁸ Bell, P. H., in Thyrocalcitonin and the C Cells, ed. S. Taylor (London: William Heinemann Ltd., in press).

⁹ Matthews, E. W., J. M. Mosley, R. H. Breed, L. Galante, A. Tse, T. V. Gudmondson, P. G. H. Byfield, and I. MacIntyre, in Parathyroid Hormone and Thyrocalcitonin (Calcitonin),

ed. R. V. Talmage and L. F. Belanger (New York: Excerpta Medica, in press).

¹⁰ Hubbard, R., and D. M. Kremen, Anal. Biochem., 12, 593 (1965).

¹¹ Hubbard, R., Biochem. Biophys. Res. Commun., 19, 679 (1965).

¹² Keutmann, H. T., and J. T. Potts, Jr., submitted to Anal. Biochem.

¹³ Keutmann, H. T., and J. T. Potts, Jr., in preparation.

¹⁴ Benson, J. V., Jr., M. J. Gordon, and J. A. Patterson, Anal. Biochem., 18, 228 (1967).

¹⁵ Hirs, C. H. W., J. Biol. Chem., 219, 611 (1956).

¹⁶ Anfinsen, C. B., and E. Haber, J. Biol. Chem., 236, 1361 (1961).

¹⁷ Ellman, G. L., Arch. Biochem. Biophys., 82, 70 (1959).

¹⁸ Edman, P., Ann. N.Y. Acad. Sci., 88, 602 (1960).

¹⁹ Edman, P., and G. Begg, Eur. J. Biochem., 1, 80 (1967).

 20 Gray, W. R., in *Methods in Enzymology*, ed. C. H. W. Hirs (New York: Academic Press, 1967), vol. 11, p. 469.

 21 Pisano, J., W. Vanden Heuvel, and E. Horning, *Biochem. Biophys. Res. Commun.*, 7, 82 (1962).

²² Benisek, W. F., and R. D. Cole, Biochem. Biophys. Res. Commun., 20, 655 (1965).

²³ Himmelhoch, S. R., and E. A. Peterson, Biochemistry, in press.

²⁴ Potts, J. T., Jr., in Methods in Enzymology, ed. C. H. W. Hirs (New York: Academic Press, 1967), vol. 11, p. 648.

²⁵ Potts, J. T., Jr., A. Berger, J. Cooke, and C. B. Anfinsen, *J. Biol. Chem.*, 237, 1851 (1962).

²⁶ Light, A., in Methods in Enzymology, ed. C. H. W. Hirs (New York: Academic Press, 1967), vol. 11, p. 417.

²⁷ Benson, J. V., R. T. Jones, J. Cormick, and J. A. Patterson, Anal. Biochem., 16, 91 (1966).