AMINO ACID SEQUENCE OF SALMON ULTIMOBRANCHIAL CALCITONIN

BY H. D. NIALL, H. T. KEUTMANN, D. H. COPP, AND J. T. POTTS, JR.

MASSACHUSETTS GENERAL HOSPITAL, BOSTON, AND UNIVERSITY OF BRITISH COLUMBIA, VANCOUVER, BRITISH COLUMBIA

Communicated by Paul C. Zamecnik, July 31, 1969

Abstract.—Salmon ultimobranchial calcitonin has been isolated and rendered pure, as demonstrated by several chemical criteria. Its amino acid sequence was determined by means of manual Edman degradation of the intact molecule and of several peptide subfragments. Results of automated degradation provided confirmation of the structure. The salmon molecule possesses, in common with other calcitonins, a 32-amino acid peptide chain terminating in prolinamide and containing half-cystine residues at positions 1 and 7. Although the sequence of the salmon hormone differs considerably from that of the porcine, bovine and human calcitonins, the four hormones are homologous in 9 of 32 positions. The much higher biological potency possessed by the salmon calcitonin makes it of particular interest for future structure function studies.

Recent interest in the physiological role of the peptide hormone calcitonin has led to the detailed study[1–4] by chemical, immunological, and biological methods of material obtained from several mammalian species. Measurements of secretion rate by bioassay[5] and radioimmunoassay[6] have confirmed the hormonal role of calcitonin in man and other mammalian species.

The covalent structures of both the porcine[7] and bovine[3, 8] calcitonins were reported earlier by our laboratory. Independent studies[9, 10] by other groups have also established the porcine structure; more recently Neher *et al.*[11] have established the amino acid sequence of the human hormone. Considerable structural differences were found between human calcitonin and the hormones from the other mammalian species. This has increased interest in the comparative chemistry of calcitonin in relation to the structural features essential for activity.

In mammals the major source of calcitonin is the thyroid gland, where it is secreted from the parafollicular or "C" cells.[12] It has been discovered by Copp *et al.*[13] that in a number of nonmammalian species (particularly birds, fish, and reptiles), calcitonin derives not from the thyroid but from a separate organ, rich in "C" cells, the ultimobranchial body. Since the ultimobranchial calcitonin is active in mammalian species, it seemed likely that the amino acid sequence of a non-mammalian calcitonin, apart from its considerable phylogenetic interest, could provide further information on the structural requirements for hypocalcemic activity of the hormone. The availability of ultimobranchial calcitonin from the salmon provided the opportunity for such a study; O'Dor *et al.*[14] and Keutmann *et al.*[15] have reported the isolation of pure salmon calcitonin and the determination of its amino acid composition. The material from this species was of particular interest because it was found to possess much greater hypocalcemic

activity by both rat and mouse bioassay than the other calcitonins so far investigated.[14, 15]

We have determined the complete amino acid sequence of salmon ultimobranchial calcitonin by an extended degradation by the phenylisothiocyanate procedure[16] on the reduced carboxymethylated hormone. The sequence found was in full agreement with other data from the amino acid composition and Edman degradation of peptide fragments of the salmon molecule prepared by chemical and enzymic cleavage. Synthetic studies by Guttmann *et al.*[17] based on the amino acid sequence obtained by us have given peptide material identical to the native hormone by multiple chemical criteria and tests of biological activity. These findings are consistent with the proposed structure.

Material and Methods.—Partially purified calcitonin obtained from salmon ultimobranchial tissue (from several *Oncorhynchus* species) by a previously described method[18] was subjected to final purification by gel filtration on Sephadex G-50 followed by ion exchange chromatography on carboxymethyl cellulose (CM-52) using a linear ammonium acetate gradient. The successive purification steps were monitored by bioassay as well as by thin-layer chromatography, disc gel electrophoresis, amino acid analysis, and end-group analysis by the phenylisothiocyanate method. The final product was found to be homogeneous by each of these criteria. Evidence was obtained of a second peptide component with hypocalcemic activity, present in lower concentrations in the extracts, which was clearly separated from the major hormone fraction by ion exchange chromatography.[15] Preliminary studies indicate that this specific material has a somewhat different amino acid composition than the major calcitonin fraction, but has similar biological activity. The nature of this minor calcitonin peptide is currently being further investigated; it may represent a species variant.

Amino acid analyses were carried out using the Beckman model 121 automatic amino acid analyzer equipped for high-sensitivity analyses and rapid elution schedules.[19, 20] Total enzymic digestion of peptides was performed using papain and aminopeptidase M, and acid hydrolyses were carried out in 5.7 *N* HCl at 110°, usually in the presence of mercaptoethanol (1:2,000 v/v).[21] Reduction and carboxymethylation was performed as previously described.[22] Acetylation was performed in the pH-stat using acetic anhydride at pH 8.0.[23] Completeness of acetylation of the epsilon-amino groups of lysine was confirmed by amino acid analysis of the acetylation product after total enzymic digestion. Tryptic digestions were carried out at a 1:125 (mole/mole) enzyme: substrate ratio using trypsin treated with Tosyl-amido-2-phenylethyl-chloromethyl-ketone (TPCK).[24] Edman degradation was carried out by both manual and automated techniques. The manual degradation is based on the earlier three-stage procedure,[26] modified to allow accelerated analysis and to minimize extractive losses of shorter and more hydrophobic peptides.[26] The coupling reaction was carried out at 50°C for 30 min in 0.5 *M* dimethylallylamine-trifluoroacetate buffer, pH 9.5, in 60% aqueous pyridine. After extraction with a small volume of benzene, and lyophilization, the thiazolinone derivative was cleaved from residual peptide by incubation for 7 min at 50°C in anhydrous trifluoroacetic acid. Extraction and drying steps have been shortened from the previously described procedure, so that a complete cycle may be completed in 90 min., and 3–5 amino acids removed and identified per day. Using these modifications a degradation of 31 cycles was possible, sufficient to provide the complete sequence of the salmon calcitonin molecule. Automated degradation was carried out by the sequenator[27] procedure modified for use with volatile reagents[8] on a Beckman-Spinco protein/peptide "Sequencer." Amino acid phenylthiohydantoin derivatives were identified by gas chromatography[26, 28] and thin-layer chromatography,[27] except for the arginine derivative which was identified by the Sakaguchi reaction.

Results.—Amino acid analysis: The amino acid composition of the purified salmon calcitonin is shown in Table 1. Tryptic digestion gave the expected four peptides (based on the presence of 2 lysines and 1 arginine), with no evidence of chymotryptic or other cleavages. Each of the peptides separated on a preparative scale by thin-layer chromatography was eluted from the plate with 50 per cent acetic acid and subjected to acid hydrolysis and amino acid analysis. The amino acid composition of each peptide is shown in Table 1. Peptide T_1 could be identified as amino terminal in location since it was the only peptide containing half-cystine, known from end-group analysis to be the amino terminal residue in the intact molecule. T_4 was felt to be the carboxyl terminal peptide since it contained no basic residue.

TABLE 1. *Acid hydrolysis (24 hr, 110°C) of native salmon calcitonin and its tryptic subfragments.**

Amino acid	Native salmon calcitonin	Tryptic Peptides, Native Calcitonin				Tryptic Peptides, Acetylated Calcitonin		Residues per mole (combined results)
		T-1	T-2	T-3	T-4	A	B	
Aspartic acid†	2.00	1.05	...	...	1.00	1.10	1.05	2
Threonine	5.10	0.95	...	1.05	2.85	1.90	3.05	5
Serine	3.50	1.90	1.15	...	1.25	2.80	1.05	4
Glutamic acid†	3.25	...	2.15	1.10	...	2.80	...	3
Proline‡	2.15	...	...	1.00	0.90	...	1.20	2
Glycine	2.90	1.05	...	...	2.10	1.10	2.00	3
Half-Cystine§	1.95	(1.10)	...	...	...	(0.55)	...	2
Valine	1.05	1.05	...	...	...	1.05	...	1
Leucine	5.05	2.05	1.75	1.15	...	5.15	...	5
Tyrosine	0.95	...	...	0.95	...	0.90	...	1
Lysine	2.05	1.00	1.00	...	...	1.95	...	2
Histidine	1.05	...	1.00	...	...	0.95	...	1
Arginine	0.95	...	...	0.95	...	1.00	...	1
							Total	32

* All values expressed as moles of amino acid per mole calcitonin.

† Two residues asparagine and two residues glutamine found on enzymatic digestion. Tryptophan absent by both enzymatic digestion and Ehrlich's stain.

‡ Proline content of peptide T-1 and acetylated peptide "A" could not be accurately determined because of coelution with reduction products of cystine formed during hydrolysis of the native peptides. Proline value for the native calcitonin was obtained using a sample subjected to performic acid oxidation prior to hydrolysis.

§ Half-cystine content of native molecule determined as cysteic acid following performic acid oxidation; cystine was partially destroyed in peptide T-1 and acetylated peptide "A," which were hydrolyzed in the native state.

Further compositional information was obtained by isolation of the two peptides obtained from a tryptic digest of acetylated calcitonin. Salmon calcitonin (5 mg) was exhaustively acetylated with acetic anhydride. This rendered the lysine residues resistant to tryptic cleavage. Hence, on tryptic digestion a single peptide-bond cleavage occurred, carboxyl terminal to the single arginine residue, giving two fragments; these were then separated from one another and from the reagents by gel filtration on Sephadex G-25 and subjected to amino acid analysis. The results are shown in Table 1; the smaller fragment was identical in composition to T_4, and the larger consisted of residues equivalent to T_1, T_2, and T_3. From the specificity of tryptic cleavage (the arginine residue must be carboxyl terminal in T_3), peptide T_3 can be placed third from the amino terminus in the

order of tryptic peptides. Since peptide T_1 was already known to be the amino terminal fragment, the order of the 4 tryptic peptides was thus established.

Edman degradation: Reduced and carboxymethylated calcitonin was subjected to 31 cycles of degradation by the accelerated manual procedure. At each step the phenylthiohydantoin derivative obtained was identified and measured quantitatively by gas chromatography. The only exception was the arginine at position 24 which was identified qualitatively by the Sakaguchi reaction. An unambiguous sequence for the first 31 amino acids was obtained, as illustrated in Figure 1. The hormone was known from its amino acid composition to contain 32 amino acids. Since 31 amino acids were identified and positioned by this

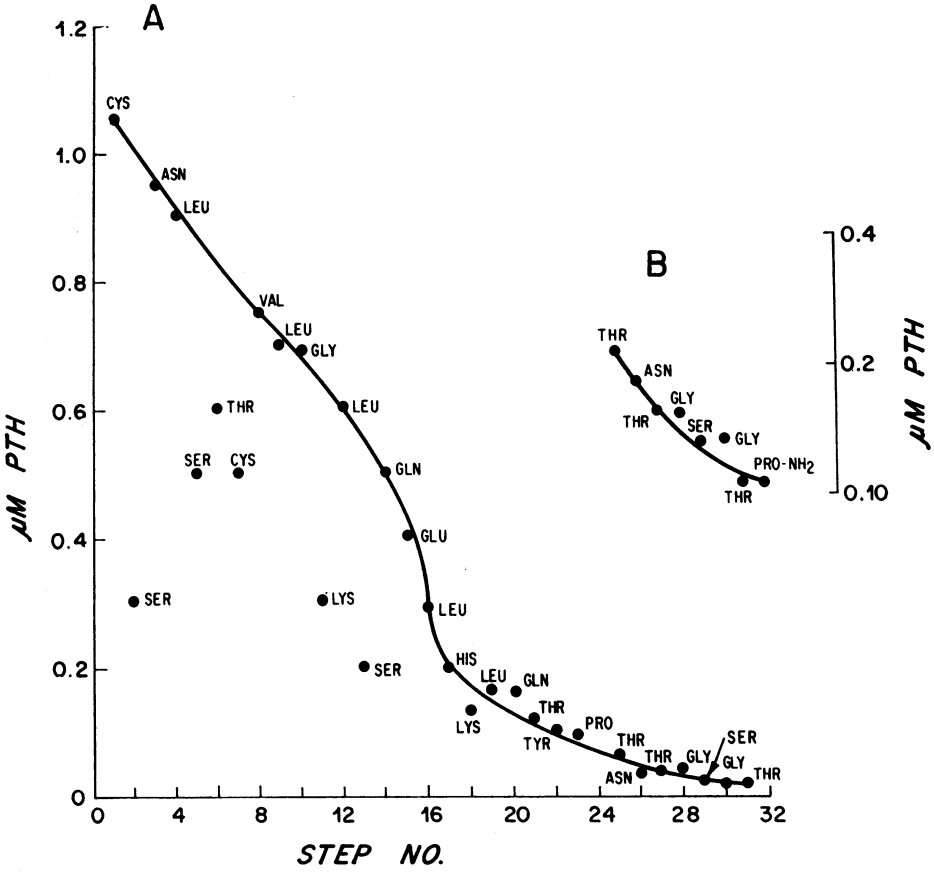

FIG. 1.—Repetitive yields of phenylthiohydantoin amino acids obtained during Edman degradation. A: degradation on 1.2 μM reduced and alkylated salmon calcitonin. Because of increasing losses of peptide material, the solvent extractions were reduced after the 16th cycle, allowing the degradation to reach the carboxyl terminus of the molecule. B: degradation on the peptide fragment (residues 25–32) obtained from tryptic digestion of acetylated salmon calcitonin. The carboxyl terminal amino acid (prolinamide) was identified directly by amino acid analysis. Recovery of cystine, lysine, threonine, and serine is lower than expected at any stage in degradation because of losses during conversion of the thiazolinone to the corresponding phenylthiohydantoin.

degradation, the only residue unaccounted for was the second of the two prolines in the molecule. This therefore must occupy position 32, the carboxyl terminal residue. The residue remaining in the reaction vessel after 31 cycles was placed directly on the basic column of the amino acid analyzer. A small peak was seen in the elution position of authentic prolinamide (the ninhydrin color value of prolinamide is relatively low). The side chain amide groups (asparagine at positions 3 and 26, glutamine at positions 14 and 20) were clearly identified by direct analysis of the appropriate amino acid amide phenylthiohydantonin.

Confirmatory Studies.—Degradation on peptide fragment (residues 25–32): In order to confirm the sequence at the carboxyl terminus of the molecule and to strengthen the identification of the carboxyl terminal amino acid as prolinamide, the fragment corresponding to T_4 (residues 25–32) was isolated by the procedure described above, and subjected to seven cycles of Edman degradation. The yields (of the phenylthiohydantoins) are shown in Figure 1. The residue after seven cycles was subjected to amino acid analysis; prolinamide was obtained in good yield. These results confirmed the sequence of residues 25–32 from the long degradation and made definite the presence of prolinamide as the carboxyl terminal residue.

Automated degradation: Apparatus for automated degradation was available only after completion of the structure by the manual technique. However, a single automated degradation was performed; this provided confirmation of the sequence of the first 25 residues. Because the extraction programs had not yet been fully developed, yields were about 25 per cent lower than those obtained with the manual method. These extractive losses prevented definitive identification after the 25th residue. We had insufficient material remaining from the supply available for sequence determination to repeat the degradation with a reduced solvent extraction program, a modification which should allow a complete degradation.

Degradation on unfractionated tryptic digest: We had earlier reported that the accelerated manual degradation is useful in the analysis of heterogeneous peptide mixtures.[25] Multiple phenylthiohydantoin (PTH) derivatives are, of course, obtained at each step. However, if the amino acid composition of each peptide in the mixture is known, it is often possible to obtain definite sequence information. The method is very economical of material since the peptides are not subjected to the losses usually encountered during purification procedures. A theoretical discussion of some potential advantages of sequence analysis on peptide mixtures has been published.[29]

In the present work it was possible to provide rapid confirmation of a major portion of the sequence by performing six cycles of Edman degradation on 2 mg of an unfractionated tryptic digest of the reduced carboxymethylated calcitonin. The following amino acids were obtained as their PTH derivatives: Step 1, cystine (as *S*-carboxymethylcystine), leucine, and threonine; Step 2, serine, glutamine, and asparagine; Step 3, threonine, glutamine, and asparagine; Step 4, leucine, glutamic acid, tyrosine, and glycine; Step 5, serine, leucine and proline; Step 6, threonine, histidine, arginine, and glycine.

The principle used in sequence analysis by this method may be illustrated by

considering the amino acids released in the 4th step. Since tyrosine occurs only in T_3, it must occupy position 4 in that peptide. Similarly, the glutamic acid must come from position 4 in T_2, the only other peptide (besides T_3) containing this residue (see Table 1). Leucine must occupy the fourth position in T_1 (T_4 contains none). By exclusion, the glycine must occupy position 4 in T_4. By similar reasoning, the amino acids occupying every position in T_2 and T_3 and the first five positions in T_1 and T_4 may be determined. Quantitation of the yields of PTH derivatives obtained was helpful in indicating release of the same amino acid from two peptides at several steps. The results were in complete agreement with the sequence obtained by the long degradation, and provided independent confirmation of a total of 23 positions.

Discussion.—The amino acid sequence of salmon calcitonin was established by the use of Edman degradation on the whole hormone molecule, and confirmed by evidence obtained from compositional analysis and Edman degradation on peptide fragments. The proposed structure receives independent support from the recent synthesis by Guttmann *et al.*[17] of fully active hormone using the sequence obtained in our studies.

Three new approaches to peptide sequence determination were used in this work. The accelerated manual procedures for Edman degradation allowed the entire sequence of 32 residues to be determined by a single degradation on five mg of calcitonin. The sequenator procedure adapted for use with peptides was most effective in the rapid confirmation of a major part of the sequence. The third method, sequence analysis by Edman degradation on peptide mixtures also provided extensive confirmatory sequence information. A total of 20 mg of calcitonin was used in all studies reported here. However, it is clear from this work and from present studies in our laboratory on mammalian calcitonins that Edman degradation by either the accelerated manual or the sequenator procedures enables complete sequence analysis to be performed on much smaller amounts of material. This could be critical in studying the hormones from other species in which there may be only a limited supply of hormone available.

Figure 2 shows the amino acid sequence of porcine, bovine, human, and salmon calcitonins. As reported by us earlier,[3, 8] the bovine structure was determined by automated Edman degradation with volatile reagents using the native molecule and fragments obtained after acetylation and tryptic digestion.

Certain consistent structural features can already be seen. The chain length is constant at 32 amino acids. The sequence for the amino terminal nine residues is identical in the porcine, bovine, and salmon molecules. The human sequence is identical in seven of these nine amino terminal positions. The sequence from all four species includes half-cystine residues at positions 1 and 7, most probably forming the disulfide bridge shown to exist in porcine[7] and human.[11] The glycine residue at position 28 and the carboxyl terminal prolinamide are also constant. The bovine and porcine structures are extremely similar, differing only in 3 positions. The human sequence more closely resembles the salmon (16 positions identical) than it does the other two mammalian hormones (14 positions identical).

Of the ten amino acids found in all four hormones, seven occupy the constant

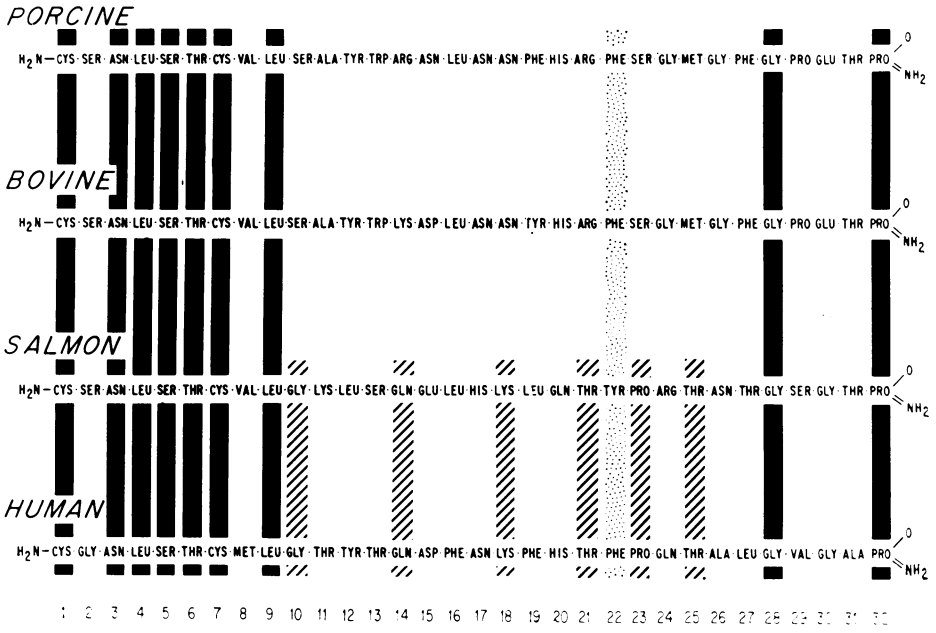

FIG. 2.—Comparison of amino acid sequences of porcine, bovine, human, and salmon calcitonins. Solid bars indicate sequence positions homologous among all four molecules. Cross-hatched bars indicate the additional positions of homology between salmon and human hormone; stippled bar indicates position where either phenylalanine or tyrosine is found in each of the calcitonins. There is also homology (not shown by cross-hatched bar) between human and salmon calcitonin at position 30.

positions enumerated above. In addition, histidine, tyrosine, and valine appear at least once in each molecule at varying positions.

Between the two extremities of the peptide chain a great deal of variation seems to be tolerated. However, some regular features are still apparent. Hydrophobic residues (tyrosine, phenylalanine, or leucine) occur at almost regular intervals along the molecule (positions 4, 9, 12, 16, 19, and 22). An aromatic residue is constantly seen at position 22. Acidic residues are confined to positions 15 and 30; basic residues are more widely distributed. Where substitutions occur for basic or acidic residues, only neutral amino acids are found. Though most substitutions are conservative, at several positions variations not explainable by single base changes in the codon are seen.

It is clear that a characteristic pattern for the covalent structure of the calcitonins can already be defined. Because of the high potency of its hypocalcemic activity, the salmon molecule will be of particular interest in future investigations of the structural basis for biological activity of calcitonins.

The authors wish to thank Miss Bess Dawson and Mr. Robert Sauer for their capable technical assistance throughout. We also gratefully acknowledge Dr. Robert Schlueter and his associates at the Armour Pharmaceutical Company, Kankakee, Illinois, for preparation of the partially purified calcitonin which made this structural analysis possible. We would like to thank Dr. D. F. Miller and the Canadian Fishing Company of Vancouver, Canada, for their cooperation in collecting the salmon ultimobranchial tissue.

This collection was facilitated by a grant to Dr. D. H. Copp from the Medical Research Council of Canada. The work described in this paper was supported by grants AM 11794 and AM 04501 from the National Institute of Arthritis and Metabolic Diseases.

[1] Munson, P. L., P. F. Hirsch, H. B. Brewer, R. A. Reisfeld, C. W. Cooper, A. B. Wasthed, H. Orimo, and J. T. Potts, Jr., *Recent Prog. Hormone Res.*, **24**, 589 (1968).

[2] Hirsch, P. F., and P. L. Munson, *Physiol. Rev.*, in press.

[3] Potts, J. T., Jr., and L. J. Deftos, Parathyroid hormone, thyrocalcitonin, vitamin D and diseases of bone and bone mineral metabolism. *In* "Duncan's Diseases of Metabolism," 6th Edition, P. K. Bondy (ed.), (Philadelphia: W. B. Saunders Co., 1969), chapter 19, pp. 904–1082.

[4] Copp, D. H., in *Mineral Metabolism*, ed. F. Bronner, (New York: Academic Press, in press), vol. 3.

[5] Gudmundsson, T. V., N. J. Y. Woodhouse, T. D. Osafo, L. Galante, E. W. Matthews, and I. MacIntyre, *Lancet*, **1**, 443 (1969).

[6] Deftos, L. J., M. R. Lee, and J. T. Potts, Jr., these PROCEEDINGS, **60**, 293 (1968).

[7] Potts, J. T., Jr., H. D. Niall, H. T. Keutmann, H. B. Brewer, and L. J. Deftos, these PROCEEDINGS, **59**, 1321 (1968).

[8] Niall, H. D., H. Penhasi, P. Gilbert, R. C. Myers, F. G. Williams, and J. T. Potts, Jr., *Fed. Proc.*, **28**, 661 (1969).

[9] Bell, P. H., W. R. Barg, Jr., D. F. Colucci, C. M. Davies, C. Dziobkowski, M. E. Englert, E. Heyder, R. Paul, and E. H. Snedeker, *J. Amer. Chem. Soc.*, **90**, 2704 (1968).

[10] Neher, R., B. Riniker, H. Zuber, W. Rittel, and F. Kahnt, *Helv. Chim. Acta*, **51**, 917 (1968).

[11] Neher, R., B. Riniker, W. Rittel, and H. Zuber, *Helv. Chim. Acta*, **51**, 1900 (1968).

[12] Foster, G. V., I. MacIntyre, and A. G. E. Pearse, *Nature*, **203**, 1029 (1964).

[13] Copp, D. H., D. W. Cockcroft, and Y. Kueh, *Canad. J. Physiol. Pharmacol.*, **45**, 1095 (1967).

[14] O'Dor, R. K., C. O. Parkes, and D. H. Copp, *Canad. J. Biochem.*, in press.

[15] Keutmann, H. T., J. A. Parsons, R. J. Schlueter, and J. T. Potts, Jr., *J. Biol. Chem.*, in press.

[16] Edman, P., *Acta Chem. Scand.*, **4**, 283 (1950).

[17] Guttmann, St., J. Pless, R. A. Huguenin, E. Sandrin, H. Bossert, and H. Willems, *Helv. Chim. Acta*, in press.

[18] O'Dor, R. K., C. O. Parkes, and D. H. Copp, *Comp. Biochem. Physiol.*, in press.

[19] Hubbard, R., and D. M. Kremen, *Anal. Biochem.*, **12**, 593 (1965).

[20] Hubbard, R., *Biochem. Biophys. Res. Commun.*, **19**, 679 (1965).

[21] Keutmann, H. T., and J. T. Potts, Jr., *Anal. Biochem.*, **29**, 175 (1969).

[22] Anfinsen, C. B., and E. Haber, *J. Biol. Chem.*, **236**, 1361 (1961).

[23] Riordan, J. F. and B. L. Vallee, in *Methods in Enzymology*, ed. C. H. W. Hirs (New York: Academic Press, 1967), vol. 11, p. 565.

[24] Kostka, V., and F. M. Carpenter, *J. Biol. Chem.*, **239**, 1799 (1964).

[25] Edman, P., *Ann. N.Y. Acad. Sci.*, **88**, 602 (1960).

[26] Niall, H. D., and J. T. Potts, Jr., Proceedings of the First American Peptide Symposium (New York: Marcel Dekker, 1969, in press).

[27] Edman, P., and G. Begg, *Eur. J. Biochem.*, **1**, 80 (1967).

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

[29] Gray, W. R., *Nature*, **220**, 1300 (1968).