Calcitonin Messenger RNA Encodes Multiple Polypeptides in a Single Precursor

Abstract. *Recombinant DNA techniques were used to analyze the structure of the messenger RNA encoding a precursor of calcitonin, a small calcium-regulating hormone of 32 amino acids. Analyses of the nucleotide sequences of cloned complementary DNA's comprising the entire coding sequence of the messenger RNA revealed that calcitonin is flanked at both its amino and carboxyl termini by peptide extensions linked to the hormone by short sequences of basic amino acids. The location of glycine next to the carboxyl terminal prolinamide of calcitonin is consistent with indications that glycine is required for the enzymatic amidation of proline to the prolinamide. During cellular bioysnthesis, calcitonin arises from a large precursor protein by cleavages at both amino and carboxyl terminal residues of the hormone. These findings raise questions concerning the regulation of these cleavages and the potential biological functions of the precursor extensions derived from these cleavages.*

In the 1960's, Copp *et al.* (*1*) identified a humoral factor in cervical venous blood that had potent hypocalcemic actions. Subsequently Hirsch *et al.* (*2*) established that the factor originated in the thyroid gland of mammals. This factor, termed calcitonin, is a small polypeptide hormone of approximately 3500 daltons that is synthesized in the parafollicular cells of the thyroid gland of mammals and in the ultimobranchial glands in lower vertebrates (*2–4*). The hormone contains 32 amino acids, including a 1–7 disulfide bridge and a carboxyl terminal prolinamide (*5*). More recent evidence indicates that calcitonin, in addition to serving as a regulator of calcium metabolism, may also act in the central nervous system in the regulation of appetite (*6*). In preliminary investigations, calcitonin-like immunoreactivity has been localized in the brain (*7*) and the pituitary gland (*8*).

Although little is known about the cellular mechanisms involved in the biosynthesis of this hormone, we (*9*) and others (*10, 11*) found, by analyses of the proteins synthesized in cell-free systems programmed by messenger RNA's (mRNA's) prepared from thyroid parafollicular cells, that the hormone is synthesized in the form of a large precursor of apparent molecular weight 15,000 (*12*). Recently, a partial sequence of this precursor has been reported (*13*). To determine the complete amino acid sequence of this precursor, we constructed and cloned, in bacteria, recombinant plasmids containing DNA's complementary to mRNA's of the parafollicular cells from a rat. We now report the nucleotide sequence of a cloned complementary DNA (cDNA) containing all of the coding sequence for a precursor of rat calcitonin. In addition to confirming the amino acid sequence of this calcitonin, as determined by Raulais *et al.* (*14*), the nucleotide sequence indicates that extensive posttranslational processing must take place during the formation of calcitonin from its precursor.

To synthesize cDNA coding for calcitonin, we isolated total polyadenylated [poly(A)] RNA from transplantable rat carcinomas of the thyroid, which are composed primarily of parafollicular cells (*15*). Approximately 50 μg of poly(A)-rich mRNA was isolated from 1 g of tumor, by extraction with phenol followed by adsorption onto oligodeoxythymidylate-cellulose (*16*). The cDNA was synthesized from mRNA with reverse transcriptase (*17*). Double-stranded DNA was prepared from the cDNA with polymerase I and inserted into the Pst I restriction endonuclease site of the plasmid pBR322 (*17*). We introduced the recombinant plasmids into *Escherichia coli* χ1776 by standard techniques (*17*) and selected clones containing calcitonin-related cDNA's by using a highly specific hybridization-selection procedure (*18*) based on the property of a cDNA containing a coding sequence for the calcitonin precursor to hybridize specifically to the calcitonin mRNA (Fig. 1). Two of the cloned cDNA's specifically selected an mRNA that directed the cell-free synthesis of the precursor. These cDNA's were further characterized by cleavages with restriction endonucleases. Analyses of the nucleotide sequences of the cDNA's, approximately 350 and 900 base pairs in length, respectively, were determined by the chemical procedures of Maxam and Gilbert (*19*) (Fig. 2).

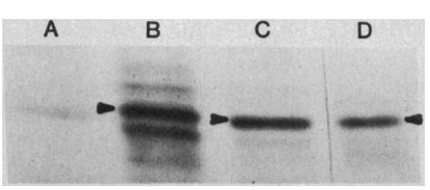

Fig. 1. Hybridization selection of calcitonin mRNA by cloned cDNA. Recombinant plasmid DNA's were cleaved with the restriction enzyme Eco RI and immobilized on nitrocellulose filters. Immobilized DNA was used to select mRNA's by hybridization (*18*). Hybridized mRNA's were eluted from the filters and translated in wheat germ cell-free assays in the presence of [^{35}S]methionine. Labeled translation products were separated on sodium dodecyl sulfate–polyacrylamide and autoradiofluorographed for 3 days at −70°C (*9*). (Lane A) Products of translation in the absence of RNA. (Lane B) Products of translation of total poly(A) RNA isolated from a rat medullary carcinoma of the thyroid. (Lane C) Products of translation of mRNA selected by a cloned cDNA. (Lane D) Immunoprecipitation of calcitonin precursor with an antiserum to human calcitonin (*9*) from total translation products in lane B. Arrows point to the calcitonin precursor with an apparent molecular weight of 15,000.

The nucleotide sequence includes portions of the 3′ and 5′ untranslated regions of the calcitonin mRNA and the entire coding region of 136 amino acids of the calcitonin precursor. The sequence coding for the 32 amino acids of calcitonin (boxed residues in Fig. 2) is located near the carboxyl terminus of the precursor. The results are in complete agreement with the structure of the hormone isolated from the thyroid gland of normal rats determined by protein sequence analysis (*14*). A dipeptide, lysylarginine, precedes and a tetrapeptide sequence, glycyllysyllysylarginine, follows the calcitonin sequence (circled residues in Fig. 2). The formation of calcitonin from its precursor appears to involve proteolytic cleavage at these basic residues by trypsin-like and carboxypeptidase B–like enzymes, similar to the cleavages of other precursors of hormones (*20*), including proinsulin (*20*), proparathyroid hormone (*21*), and the large common precursor for adrenocorticotropic hormone (ACTH), melanocyte-stimulating hormone (MSH), and the endorphins (*22*).

The nucleotide sequence indicates the presence of a glycine (hexagon-enclosed residue in Fig. 2) adjacent to the carboxyl terminal proline of calcitonin. Recent investigations of the biosynthesis of melittin, a small peptide hormone, suggest that the enzymatic amidation of carboxyl terminal amino acids of peptides requires an adjacent glycine; the glycine appears to serve as a donor of the amide group (*23*). Whether this mechanism is universal for the generation of all peptides that bear carboxyl terminal amides is unknown and must await the structural analysis of biosynthetic precursors of other proteins containing carboxyl terminal amides.

The nucleotide sequence predicts not only the structure of calcitonin, but also that of cryptic peptides contained within the precursor of calcitonin. A peptide of

0036-8075/81/0724-0457$00.50/0

16 amino acids lies at the carboxyl terminus of the precursor. In the amino terminal direction, we encountered a sequence of 84 amino acids: approximately 60 residues comprising a highly acidic cryptic peptide, preceded by a hydrophobic leader (or signal) sequence (*20, 22, 24*), and the nucleotide sequence ATG corresponding to the methionine codon AUG that signals the start of protein synthesis (Fig. 2). Approximately five bases from the start codon in the 5′ direction was a nucleotide sequence, AGGGAGG, reminiscent of a Shine-Dalgarno sequence believed to be necessary for the binding of the ribosome during translation (*25*).

The sequenced cDNA was used as a probe to explore the number and sizes of mRNA's coding for calcitonin-related proteins in the rat thyroid tumor. Analyses, by agarose-gel electrophoresis and hybridization with ^{32}P-labeled cDNA, of RNA extracted from the tumor (*26*) revealed a predominant mRNA of approximately 1000 nucleotides. Inasmuch as 408 bases of the mRNA code for the calcitonin precursor, approximately 600 nucleotides of the mRNA must consist of untranslated 3′ and 5′ regions.

From our data, we can postulate the cellular processes involved in the formation of calcitonin, 3500 daltons, from its larger precursor of 15,000 daltons: (i) cotranslational glycosylation (*27*) and cleavage of a leader sequence from the calcitonin precursor; (ii) trypsin-like cleavage of procalcitonin; (iii) subsequent trimming of basic amino acids from the carboxyl terminus of the cleaved products by carboxypeptidase B–like activity; and, finally, (iv) enzymatic conversion of the Pro-Gly (Pro, proline; Gly, glycine) amino acid sequence to the carboxyl terminal prolinamide found in the mature form of the hormone. The carbohydrate attached to the precursor might be involved in the determination of the specificity of the posttranslational cleavages that occur as the precursor migrates through the secretory pathway of the cell.

The calcitonin precursor may be one of a class of polypeptide hormone precursors that includes the common precursor for ACTH, MSH, and the endorphins (*22*) and the common precursor that contains the sequences of neurophysin and vasopressin (*28*). In these instances, two or more biologically active peptides that are secreted from the cell are specifically processed from larger, glycosylated precursors. The peptides flanking calcitonin within the precursor might have biological activities as yet unrecognized and be secreted along with calcitonin from the C cells of the thyroid. A preliminary search of peptide sequences with known functions did not

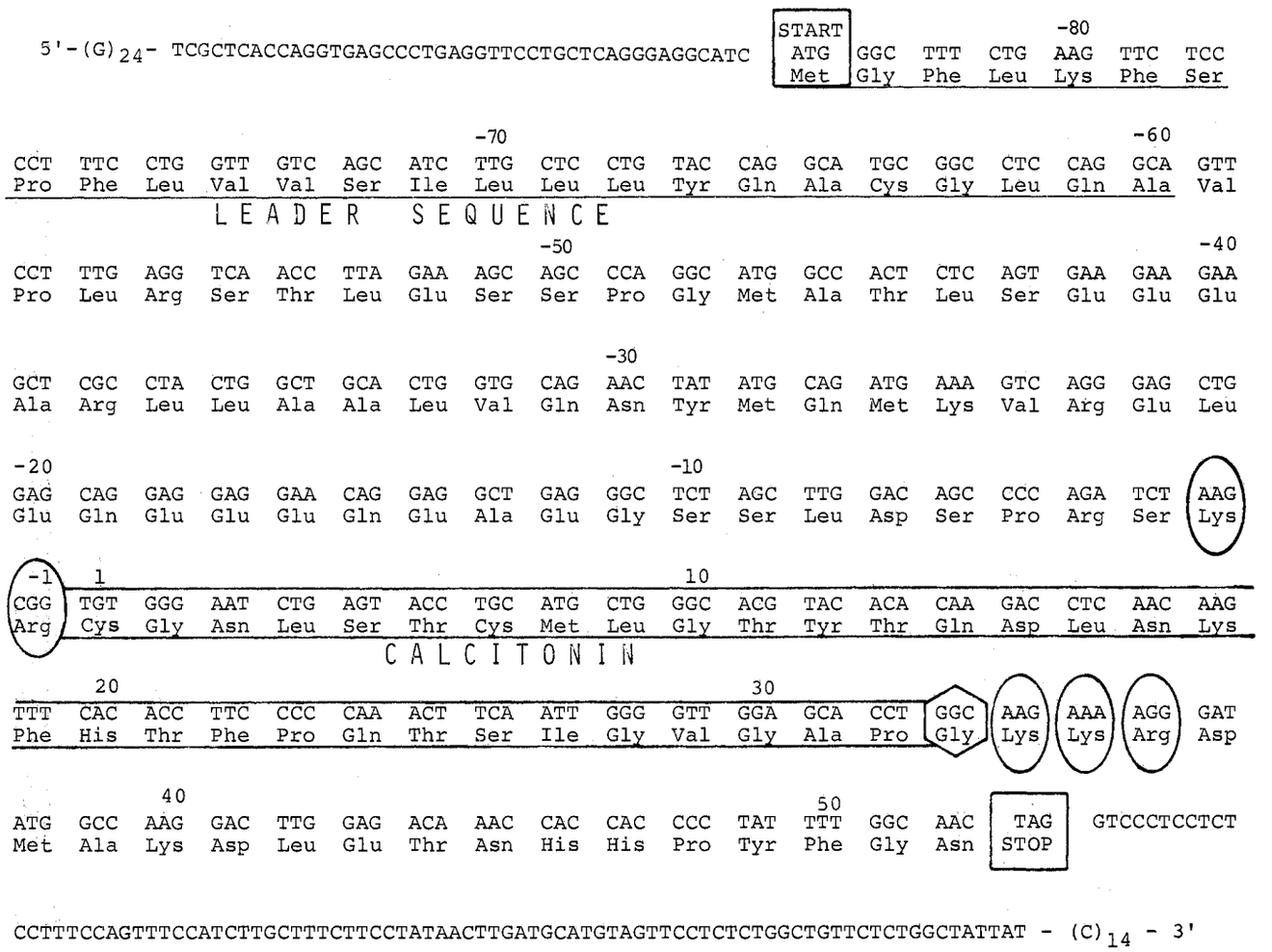

Fig. 2. Nucleotide sequence of cloned rat calcitonin cDNA. Single end-labeled DNA fragments, generated by cleavages with restriction endonucleases, were isolated by electrophoresis on polyacrylamide gels and both sense and nonsense strands were sequenced by use of base-specific, limited chemical cleavages (*19*). The sequence shown consists of 545 nucleotides of the largest cDNA, consisting of approximately 900 nucleotides. The deduced amino acid sequence of the precursor begins with methionine at position −84 and ends with asparagine at position 52. The calcitonin sequence begins at residue 1 (cysteine) and continues to residue 32 (proline). The basic amino acids flanking calcitonin are enclosed in ovals. The glycine residue at position 33 is presumed to be required for the enzymatic amidation of proline at residue 32. The initiation codon (ATG, Met) and the termination codon (TAG, stop), are enclosed in the boxes.

reveal any significant homologies between these cryptic peptides and peptides of known structure (*29*). Alternatively, these peptides may serve structural roles, important for the proper folding of the precursor into a conformation that allows its accurate proteolytic processing by tissue-specific proteases during the formation of calcitonin, or they may serve simply as inert protein spacers without specific biological activities. These cellular cleavages of the calcitonin precursor may differ in different tissues, such as the thyroid and brain, comparable with the processing of the common precursor of ACTH, MSH, and the endorphins (*30*), in which the cleavage in the anterior and in the intermediate lobe of the pituitary is different. The knowledge that the precursor is glycosylated (*29*) and contains peptide sequences attached to calcitonin raises the possibility that forms of the hormone detected in the circulation, with apparent molecular weights larger than calcitonin, may result from incomplete processing of the precursor prior to secretion in vivo (*31*). The chemical synthesis of the peptide extensions of the calcitonin precursor should be useful for several purposes, including the preparation of radioimmunoassays that can be used for the detection of the precursor-specific peptides during intracellular processing and for analyses of the immunochemical composition of the multiple circulating forms of calcitonin.

J. W. JACOBS
R. H. GOODMAN
W. W. CHIN
P. C. DEE
J. F. HABENER

Laboratory of Molecular Endocrinology, Massachusetts General Hospital, and *Howard Hughes Medical Institute Laboratories, Harvard Medical School, Boston 02114*

N. H. BELL

Medical University of South Carolina, and *Veterans Administration Medical Center, Charleston 29403*

J. T. POTTS, JR.,

Department of Medicine, Harvard Medical School, and *Endocrine Unit, Massachusetts General Hospital*

References and Notes

1. D. H. Copp, E. C. Cameron, B. A. Cheney, A. G. F. Davidson, K. G. Henze, *Endocrinology* **70**, 638 (1962).
2. P. F. Hirsch, G. F. Gauthier, P. L. Munson, *ibid.* **73**, 244 (1963).
3. G. V. Foster *et al.*, *Nature (London)* **202**, 1303 (1964); A. D. Care, *ibid.* **205**, 1289 (1965).
4. D. H. Copp, D. W. Cockcroft, Y. Kueh, *Science* **158**, 924 (1967); S. D. Tauber, *Proc. Natl. Acad. Sci. U.S.A.* **58**, 1684 (1965).
5. P. L. Munson, P. F. Hirsch, H. B. Brewer, Jr., R. A. Reisfeld, C. W. Cooper, H. Orimo, J. T. Potts, Jr., *Recent Prog. Horm. Res.* **24**, 589 (1967); J. T. Potts, Jr., H. T. Keutmann, H. D. Niall, G. W. Tregear, *Vitam. Horm.* **29**, 41 (1971).
6. W. J. Freed, M. J. Perlow, R. J. Wyatt, *Science* **206**, 850 (1979).
7. I. MacIntyre, personal communication.
8. L. J. Deftos *et al.*, *J. Clin. Endocrinol. Metab.* **47**, 457 (1978).
9. J. W. Jacobs, J. T. Potts, Jr., N. H. Bell, J. F. Habener, *J. Biol. Chem.* **254**, 10600 (1979).
10. S. G. Amara, M. G. Rosenfeld, R. S. Birnbaum, B. A. Roos, *ibid.* **225**, 2645 (1980).
11. C. J. M. Lips, J. Van der Sluys Veer, J. A. Van der Donk, R. H. Van Dam, W. H. L. Hackeng, *Lancet* **1978-I**, 16 (1978).
12. Apparent molecular weight refers to estimated sizes of pre-procalcitonin as determined by sodium dodecyl sulfate–polyacrylamide gel electrophoresis. There is a discrepancy concerning the apparent molecular weight of the calcitonin precursor. We (*9*) and Amara *et al.* (*10*) detect precursors of similar molecular weight, 15,000 and 17,000 daltons, respectively. Lips *et al.* (*11*), however, detect a much larger precursor of 60,000 daltons.
13. S. G. Amara, D. N. David, M. G. Rosenfeld, B. A. Roos, R. M. Evans, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 444 (1980).
14. D. Raulais, J. Hagaman, D.A. Ontjes, R. L. Lundblad, H. S. Kingdon, *Eur. J. Biochem.* **64**, 607 (1976).
15. B. A. Boorman and C. F. Hollander, *Am. J. Pathol.* **83**, 237 (1976); R. Pantzer, S. F. Queener, N. H. Bell, *Clin. Res.* **26**, 117 (1978); B. A. Roos, M. J. Yoon, A. L. Frelinger, A. E. Pensky, R. S. Birnbaum, P. W. Lambert, *Endocrinology* **105**, 27 (1979).
16. J. A. Majzoub, H. M. Kronenberg, J. T. Potts, Jr., A. Rich, J. F. Habener, *J. Biol. Chem.* **254**, 7449 (1979).
17. H. M. Kronenberg, B. M. McDevitt, J. A. Majzoub, J. Nathans, P. E. Sharp, J. T. Potts, Jr., A. Rich, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 4981 (1979); G. N. Hendy and H. M. Kronenberg, personal communication.
18. R. P. Ricciardi, J. S. Miller, B. E. Roberts, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 4927 (1979).
19. A. Maxam and W. Gilbert, *ibid.* **74**, 560 (1977).
20. D. F. Steiner, P. S. Quinn, S. J. Chan, J. Marsh, H. S. Tager, *Ann. N.Y. Acad. Sci.* **343**, 1 (1980).
21. J. F. Habener, H. T. Chang, J. T. Potts, Jr., *Biochemistry* **16**, 3910 (1977); D. Goltzman, E. N. Callahan, G. W. Tregear, J. T. Potts, Jr., *ibid.* **15**, 5076 (1976).
22. B. A. Eipper and R. E. Mains, *J. Biol. Chem.* **253**, 5732 (1978); S. Nakanishi, A. Inoue, T. Keta, M. Nakamura, A. C. Y. Chang, S. Cohen, S. Numa, *Nature (London)* **278**, 423 (1979).
23. G. Kreil, G. Suchanek, I. Kindås-Mügge, *Fed. Proc. Fed. Am. Soc. Exp. Biol.* **36**, 2081 (1977).
24. J. F. Habener and J. T. Potts, Jr., *N. Engl. J. Med.* **299**, 580 (1978).
25. J. Shine and L. Dalgarno, *Proc. Natl. Acad. Sci. U.S.A.* **71**, 1342 (1974).
26. J. C. Alwine, D. J. Kemp, G. R. Stark, *ibid.* **74**, 5350 (1977).
27. J. W. Jacobs, P. K. Lund, N. H. Bell, J. T. Potts, Jr., J. F. Habener, *J. Biol. Chem.* **256**, 2803 (1981).
28. J. T. Russell, M. J. Brownstein, H. Gainer, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 6086 (1979).
29. A search of 160,000 hexadecapeptide sequences in proteins studied by M. O. Dayhoff, National Biomedical Research Foundation, Washington, D.C., failed to uncover any homologies.
30. J. L. Roberts, M. A. Phillips, P. A. Rosa, E. Herbert, *Biochemistry* **17**, 3609 (1978).
31. F. R. Singer and J. F. Habener, *Biochem. Biophys. Res. Commn.* **61**, 710 (1974); L. J. Deftos, B. A. Roos, D. Bronzert, J. G. Parthemore, *J. Clin. Endocrinol. Metab.* **40**, 409 (1975); D. Goltzman and A. S. Tischler, *J. Clin. Invest.* **61**, 449 (1978); H. Heath and G. W. Sizemore, *J. Lab. Clin. Med.* **93**, 390 (1979).
32. We thank H. Kronenberg and G. Hendy for advice on the cloning experiments and J. Sullivan for manuscript preparation. Supported in part by NSF grant 79-22896 and NIH grant AM 27781-01.

10 December 1980; revised 26 March 1981

Species-Typical Behavior of Hamsters Deprived from Birth of the Neocortex

Abstract. *Hamsters deprived from birth of the neocortex developed normally and displayed the usual hamster-typical behavioral patterns. With the additional concurrent destruction of midline limbic convolutions (cingulate and underlying dorsal hippocampal), there were deficits in maternal behavior and a lack of development of play behavior. These findings demonstrate in a rodent (i) that the striatal complex and limbic system, along with the remaining neuraxis, are sufficient for giving expression to a wide range of unlearned forms of species-typical behavior and (ii) that midline limbic structures are required for the expression of play behavior and the integrated performance of maternal behavior.*

The purpose of this study was to test the hypothesis that in rodents the two older evolutionary formations of the forebrain [the striatal complex and the limbic system (*1*, *2*)] are sufficient, along with the remaining nervous system, for the expression of most forms of species-typical behavior. Respective parts of the mammalian striatal complex and limbic system reflect an ancestral relationship to reptiles and early mammals (*3*). We found that hamsters deprived from birth of the neocortex grew normally and displayed the usual hamster-typical behavioral patterns seen under laboratory conditions (*4*). The findings are at variance with the popular view that in mammals generally the neocortex is required for fully integrated, directed behavior.

The subjects were Syrian golden hamsters (*Mesocricetus auratus*) descended from wild hamsters captured in 1971 (*5*). The neocortex was eliminated by heat applied to the skull or by aspiration (under cryoanesthesia) on the first or second day after birth. Littermates used as controls received similar treatment, but without the destruction of brain tissue (Fig. 1A). After surgery the pups were returned to the dam. The day-to-day behavioral development was recorded on a checklist derived from an extensive ethogram (*6*). Time-lapse television was used for 24-hour observations of animals living in special habitats. Quantitative measures of mating and certain other behaviors were obtained by the use of a computer-assisted event recorder.

0036-8075/81/0724-0459$00.50/0