Volume 223, number 1, 63–68 FEB 05193 October 1987

Sequence and expression of the chicken calcitonin gene

S. Minvielle, M. Cressent, M.C. Delehaye, N. Segond, G. Milhaud, A. Jullienne, M.S. Moukhtar and F. Lasmoles

U 113 INSERM and UA 163 CNRS, CHU St. Antoine, 27, Rue de Chaligny, 75571 Paris Cedex 12, France

Received 28 July 1987; revised version received 26 August 1987

The avian calcitonin gene was isolated and sequenced; two mRNAs are expressed by tissue-specific alternate splicing. The peptides encoded by the mRNAs are the protein precursors of either calcitonin or calcitonin gene-related peptide (CGRP). Calcitonin is expressed predominantly in ultimobranchial bodies and CGRP in brain.

Calcitonin gene; CGRP gene; RNA processing; Hormone precursor; (Chicken ultimobranchial gland)

1. INTRODUCTION

In two mammals, man and rat, the calcitonin gene [1,2] produces two mRNAs by tissue-specific alternate splicing [3], one coding for the precursors of calcitonin [4,5] and the other for the calcitonin gene-related peptide (CGRP) [3]. Calcitonin mRNA is expressed predominantly in the thyroidal 'C' cells and the CGRP mRNA is found in the central and peripheral nervous system. Calcitonin is synthesized in the ultimobranchial glands of non-mammalian vertebrates [6].

We report here the sequence of a non-mammalian (chicken) calcitonin/CGRP gene. The gene can generate two mRNAs encoding the precursors of calcitonin and CGRP by tissue-specific alternate splicing.

2. MATERIALS AND METHODS

2.1. *Screening of the chicken genomic library, subcloning and DNA sequence analysis*

Isolation of the genomic phage λ L47.1 containing the calcitonin gene region has been described [7]. DNA fragments of interest were subcloned into pUC8 and pUC9 and used for restriction mapping.

Two large overlapping fragments were subcloned in phages M13 mp11 and mp19 [8]. A set of sequences of overlapping clones was obtained [9]. Nucleotide sequences were determined by the dideoxynucleotide chain-termination method [10].

Correspondence address: S. Minvielle, U 113 INSERM and UA 163 CNRS, CHU St. Antoine, 27, Rue de Chaligny, 75571 Paris Cedex 12, France

2.2. *Primer synthesis and extension*

A single-stranded, synthetic oligonucleotide complementary to the 5′-end of exon 2, 5′-[CACACAACCAAGGCATAAAC]-3′, was labeled at its 5′-end with [γ-^{32}P]ATP using T_4 polynucleotide kinase (BRL). Hybridization and primer extension were according to Williams and Mason [11]. The cDNA produced was subjected to electrophoresis on 8% urea-polyacrylamide gels and visualized by autoradiography.

2.3. *Northern blot analysis*

RNAs were extracted from chicken ultimobranchial glands and brain with phenol-chloroform [12], purified by LiCl precipitation and poly(A)$^+$ RNA separated on oligo(dT)-cellulose. Glyoxal-denatured RNAs were separated by electrophoresis on 1% agarose [13]. RNA fragments were transferred to GeneScreen membranes (New England

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/87/$3.50

Nuclear) and hybridized with a nick-translated probe specific to chicken calcitonin (spec. act. 2×10^8 cpm/μg). Filters were washed twice in 2 × SSC for 5 min at room temperature, followed by two washes in 2 × SSC and 1% SDS for 30 min at 55°C, two washes in 0.1 × SSC for 30 min at room temperature and exposed at −80°C to Royal X-Omat AR5 film (Kodak). Filters were dehybridized and tested with an Sp6 RNA [14] probe specific to CGRP [15] (spec. act. $>10^8$ cpm/μg). They were washed four times in 50 mM NaCl, 20 mM Na_2HPO_4, pH 6.5, 1 mM EDTA, 0.1% SDS for 20 min at 60°C and exposed as above.

3. RESULTS

3.1. *Characterization of the chicken calcitonin gene*

DNA obtained from the phage showing the highest signal with the chicken calcitonin probe was subjected to restriction analysis. Fragments containing exons were selected by hybridization with the specific probe. A 6 kb *Eco*RI-*Xba*I fragment that hybridized to the specific cDNA probe and a 3.75 kb *Xba*I-*Xba*I fragment were sequenced. The restriction map and exon-intron structure of the chicken calcitonin gene are reported in fig.1. The nucleotide sequence and the predicted amino acid sequence are shown in fig.2. Alignment of this gene with the chicken calcitonin cDNA [16], murine [1] and human [2] genes reveals that it is divided into 6 exons. All intron/exon boundaries known agree with the splice junction consensus [17].

The technique of primer extension has been used to determine the length of exon 1 (fig.3). One transcript of 141 nucleotides in length was produced; we conclude that exon 1 is 78 nucleotides long, the location of the beginning of exon 2 being inferred from the high homology with human and rat genomic sequences [1,2]. Exons 2 and 3 encode the common N-terminal sequences of the precursor of calcitonin and CGRP peptides. A single difference is observed between cDNA sequence specific for chicken calcitonin mRNA that we have previously reported [16] and the corresponding genomic sequence. This base substitution results in an amino acid change from Glu to Asp and alters the restriction site of *Eco*RI. Exon 4 encodes the calcitonin peptide; the nucleotide sequence confirms that the polyadenylation signal AATAAA is located 326 nucleotides downstream from the calcitonin stop codon. Exon 5 encodes the CGRP peptide, the 3′-end of the last exon being deduced from its striking homology with the murine and human sequences at the recognition sequence AATTAAA. Extensive homology of the chicken and mammalian genes is present in the polyadenylation region of calcitonin (fig.4A) and CGRP (fig.4B). In the calcitonin polyadenylation region, G/T clusters are present downstream from the poly(A) addition site and from the consensus sequence CATTG [18].

3.2. *Expression of the calcitonin/CGRP gene in ultimobranchial glands and brain*

In order to verify whether tissue-specific alternate switching of the calcitonin gene also occurred

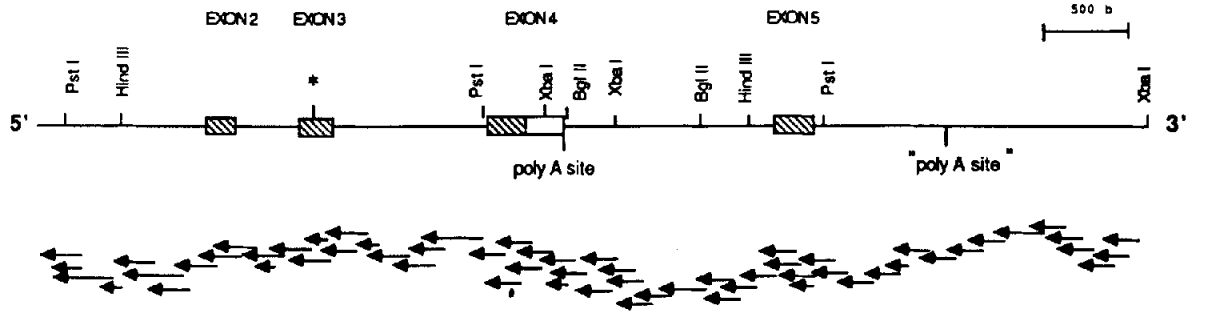

Fig.1. Restriction map and sequencing strategy for the chicken calcitonin/CGRP gene. The boxes represent exons; shaded areas in exons indicate the coding sequences. The length and direction of sequencing by the dideoxy chain termination method are shown by horizontal arrows. DNA sequencing was started at a site that was generated by sequential digestion employing the 3′ to 5′ specific exonuclease activity of T_4 DNA polymerase. The asterisk indicates the *Eco*RI site present in the cDNA but absent in the chicken calcitonin/CGRP gene.

Fig.2. Nucleotide sequence and deduced amino acid sequence of the chicken calcitonin/CGRP gene. The polyadenylation signal in the 3'-region of exon 4 and the presumptive polyadenylation signal CGRP messenger are underlined. The sequence complementary to the oligonucleotide primer P1 is indicated.

in chicken, total and poly(A)$^+$ RNAs were extracted from the ultimobranchial glands and the brain of chicken and subjected to Northern blot analysis (fig.5). Poly(A)$^+$ RNA of approx. 1.5 kb extracted from brain tissue hybridized strongly to the CGRP-specific probe. Ultimobranchial poly(A)$^+$ RNA showed a weak hybridization signal with this specific probe whereas with the calcitonin-specific probe the reverse was observed, i.e. a strong hybridization with the 1.5 kb ultimobranchial poly(A)$^+$ RNA species and a weak signal with the brain poly(A)$^+$ RNA species of the same length. The 5.1 kb long total RNA extracted from brain or ultimobranchial glands showed a strong

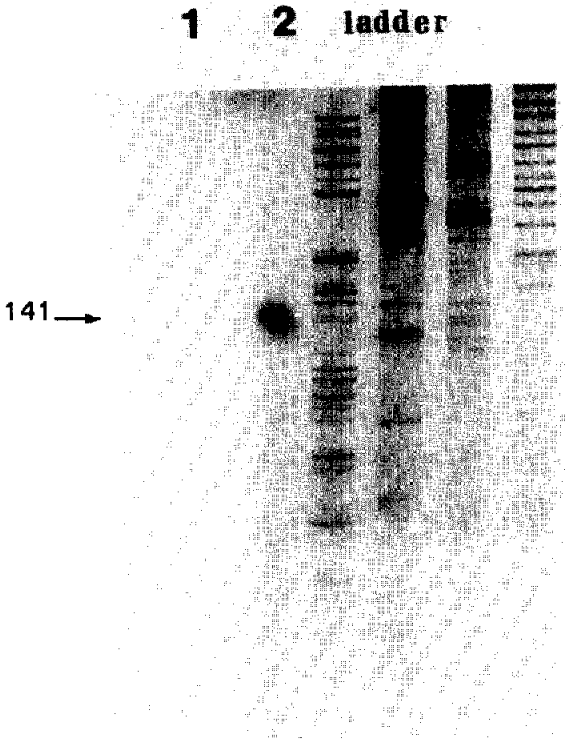

Fig.3. Determination of the length of non-coding exon 1 by primer extension. Lane 1, control reaction where RNA was omitted. Lane 2, a synthetic oligonucleotide was used to prime synthesis with reverse transcriptase from chicken ultimobranchial glands poly(A)$^+$ (10 μg). The M13 sequencing ladder of known sequence was used as a sizing standard. The size of the band indicated is in nucleotides. The primer is 20 nucleotides long and is extended for 121 nucleotides.

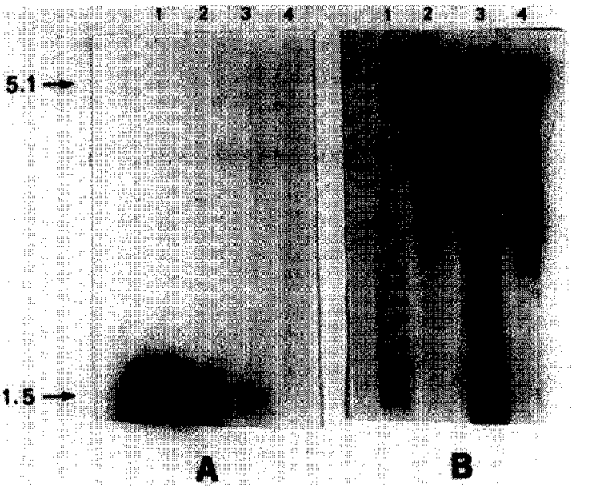

Fig.5. Northern blot analysis of total and poly(A)$^+$ RNAs. Total RNAs from chicken ultimobranchial glands and brain were separated by electrophoresis on 1% agarose and RNA fragments transferred to GeneScreen membranes were hybridized with a ^{32}P-labelled *DdeI/XbaI* restriction fragment of pCCT22 specific to the coding sequence of chicken calcitonin (A). The same filter after washing was reprobed with a ^{32}P-labelled probe specific to CGRP (B). Lanes: 1, 5 μg poly(A)$^+$ from chicken UB glands; 2, 10 μg total RNA from chicken UB glands; 3, 10 μg poly(A)$^+$ from chicken brain; 4, 10 μg total RNA from chicken brain. Size estimates are given in kb.

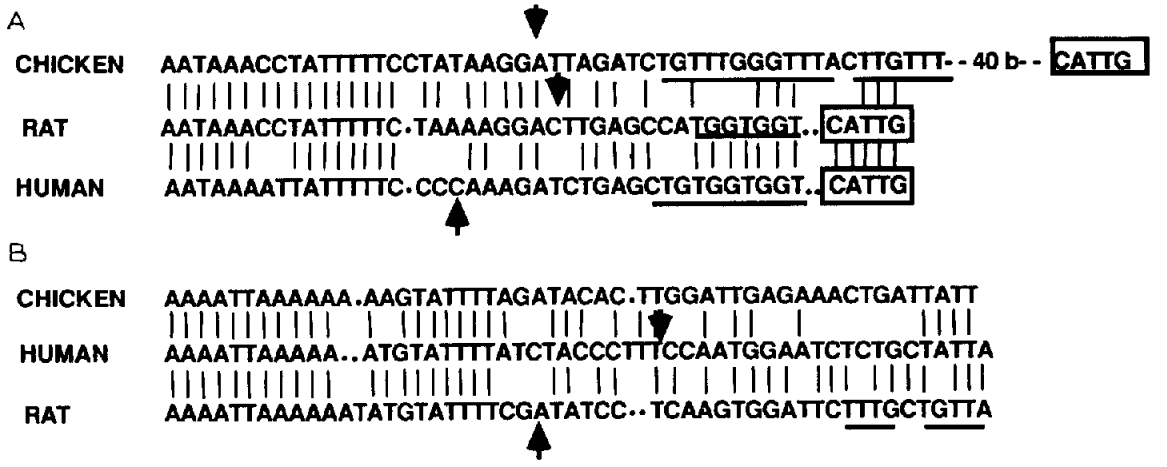

Fig.4. Comparison of sequences downstream from the first polyadenylation signal (A) and the second polyadenylation signal (B) in chicken, rat and human. G/T clusters downstream from the poly(A) addition site are underlined, and the consensus sequences CATTG are boxed. Arrowheads identify the poly(A) addition sites. Black points indicate gaps introduced to optimize the homologies.

signal with the CGRP-specific probe, but not with the calcitonin-specific probe.

4. DISCUSSION

We have used a cDNA probe specific for chicken calcitonin mRNA to isolate the calcitonin/CGRP gene from a chicken genomic library. We have elucidated the complete sequence of the protein precursors of chicken calcitonin and CGRP. The gene shows the same organization as the mammalian genes. The nucleotide sequence of exon 4, which codes for calcitonin in this gene, is identical to the cDNA sequence of chicken calcitonin, and thus fully confirms the amino acid sequence deduced from the mRNA sequence. The deduced amino acid sequence of chicken calcitonin is also identical to the sequence of chicken calcitonin recently reported by Homma et al. [19]. We observed a single difference between the genomic and mRNA sequences, i.e. the absence in the genomic sequence of an *Eco*RI site present in the messenger. This difference could be due to polymorphism of this gene; Southern blot analysis of the chicken DNA extracted from several different animals of the same strain revealed that some of them had calcitonin/CGRP genes which lacked the *Eco*RI site. It is also highly probable that two calcitonin/CGRP genes exist in the chicken, inasmuch as a second calcitonin molecule was reported in chicken [19,20] and a second CGRP molecule exists in man and rat [21,22].

Our hybridization data show that the specific mRNA for calcitonin is expressed predominantly in the ultimobranchial bodies and CGRP messenger principally in the nervous system. However, the calcitonin-specific probe did not detect RNA species longer than 1.5 kb. RNA species of 3.4 kb [3,15] are detected with calcitonin-specific probes in malignant (human cell lines or murine tumours) tissues but not in normal murine thyroid [3], where no RNA species larger than 1 kb hybridized with the calcitonin-specific probe. Thus, two distinct RNA species coexist in the chicken ultimobranchial bodies and in the brain: one species of 1.5 kb containing the calcitonin exon and another of 5.1 kb containing the CGRP exon. The larger species in the ultimobranchial bodies did not seem to be further spliced into the mature form of the CGRP messenger in the ultimobranchial bodies. The 1.5 kb mature form of both calcitonin and CGRP messengers was predominantly expressed in the ultimobranchial bodies (calcitonin) and in the brain (CGRP).

The nucleotide sequence of the introns of the calcitonin gene is of particular interest as this gene contains two distinct polyadenylation sites and its transcripts undergo alternate splicing. The regions directly upstream or downstream from the poly(A) addition sites are highly conserved in the calcitonin/CGRP gene. However, in contrast to the calcitonin polyadenylation site where G/T clusters downstream of the poly(A) addition site and consensus sequence CATTG [18] are present, the CGRP polyadenylation site is devoid of such features but contains the sequence AATTAAA which is reported to decrease 3′-maturation in vitro [23]. These results suggest that the two polyadenylation sites have different reactivity.

Thus, both splicing and polyadenylation processing of the two exons of the calcitonin gene are associated with distinct nucleotide sequences. So the putative trans-acting factors implicated in the splicing and polyadenylation could act on well-differentiated consensus nucleotide sequences in the introns surrounding these exons.

The alternate splicing of the calcitonin/CGRP gene is now being actively investigated in mammals [24,25]. The present work opens the way to the extension of such studies to non-mammalian vertebrates, a prerequisite for studying the evolution of the tissue-specific splicing mechanism.

ACKNOWLEDGEMENTS

We thank Dr M. Ballivet for making available the chicken genomic library and Professor R.K. Craig for the CGRP-specific probe. We are deeply indebted to Professor F. Chapeville for his comments and helpful criticism. This work was supported in part by a grant from L'Association pour la Recherche sur le Cancer. S.M. and M.C.D. are recipients of grants from the Ministère de la Recherche et de la Technologie.

REFERENCES

[1] Amara, S.G., Evans, R.M. and Rosenfeld, M.G. (1984) Mol. Cell. Biol. 4, 2151–2160.

[2] Jonas, V., Lin, C.R., Kawashima, E., Semon, D., Swanson, L.W., Mermod, J.J., Evans, R.M. and Rosenfeld, M.G. (1985) Proc. Natl. Acad. Sci. USA 82, 1994–1998.
[3] Amara, S.G., Jonas, V., Rosenfeld, M.G., Ong, E.S. and Evans, R.M. (1982) Nature 298, 240–244.
[4] Copp, D.H., Cameron, E.C., Cheney, B.A., Davidson, A.G.F. and Henze, K.G. (1962) Endocrinology 70, 638–649.
[5] Hirsch, P.F., Gauthier, G.F. and Munson, P.L. (1963) Endocrinology 73, 244–252.
[6] Copp, D.H., Cockcroft, D.W. and Kueh, Y. (1967) Science 158, 924.
[7] Minvielle, S., Cressent, M., Lasmoles, F., Jullienne, A., Milhaud, G. and Moukhtar, M.S. (1986) FEBS Lett. 203, 7–10.
[8] Messing, J., Crea, R. and Seeburg, P.H. (1981) Nucleic Acids Res. 9, 309–321.
[9] Dale, R.M.K., McClure, B.A. and Houchins, J.P. (1985) Plasmid 13, 31–40.
[10] Sanger, F., Nicklens, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[11] Williams, J.G. and Mason, P.J. (1985) in: Nucleic acid hybridisation - A Practical Approach (Hames, B.D. and Higgins, S.J. eds) pp. 152–160, IRL Press, Oxford.
[12] Itoh, N., Nose, K. and Okamoto, H. (1979) Eur. J. Biochem. 97, 1–9.
[13] Thomas, P.S. (1980) Proc. Natl. Acad. Sci. USA 77, 5201–5205.
[14] Melton, D.A., Kreig, P.A., Rebagliati, M.R., Maniatis, T., Zinn, K. and MacGreen, M.R. (1984) Cell 34, 865–879.
[15] Edbrooke, M.R., Parker, D., McVey, J.H., Riley, J.H., Sorenson, G.D., Pettengill, O.S. and Craig, R.K. (1985) EMBO J. 4, 715–724.
[16] Lasmoles, F., Jullienne, A., Day, F., Minvielle, S., Milhaud, G. and Moukhtar, M.S. (1985) EMBO J. 4, 2603–2607.
[17] Breatnach, R. and Chambon, P. (1981) Annu. Rev. Biochem. 50, 349–383.
[18] Berget, S.M. (1984) Nature 309, 179–182.
[19] Homma, T., Watanabe, M., Hirose, S., Kanai, A., Kangawa, K. and Matsuo, H. (1986) J. Biochem. 100, 459–467.
[20] Nieto, A., Moya, F. and R-Candela, J.L. (1973) Biochim. Biophys. Acta 322, 383–391.
[21] Steenbergh, P.H., Hoppener, J.W.M., Zandberg, J., Lips, C.J.M. and Jansz, H.S. (1985) FEBS Lett. 183, 403–407.
[22] Amara, S.G., Arriza, J.L., Leff, S.E., Swanson, L.W., Evans, R.M. and Rosenfeld, M.G. (1985) Science 229, 1094–1097.
[23] Birnstiel, M.L., Busslinger, M. and Strub, K. (1985) Cell 41, 349–359.
[24] Leff, S.E., Evans, R.M. and Rosenfeld, M.G. (1987) Cell 48, 517–524.
[25] Crenshaw, E.B. iii, Russo, A.F., Swanson, L.W. and Rosenfeld, M.G. (1987) Cell 49, 389–398.