A second human calcitonin/CGRP gene

P.H. Steenbergh, J.W.M. Höppener, J. Zandberg*, C.J.M. Lips+ and H.S. Jansz*

*Institute of Molecular Biology, State University of Utrecht, 8 Padualaan, 3584 CH Utrecht, *Laboratory for Physiological Chemistry, State University of Utrecht, 24a Vondellaan, 3521 GG Utrecht,* and *+Department of Internal Medicine, University Hospital, State University of Utrecht, 101 Catharijnesingel, 3511 GV Utrecht, The Netherlands*

Received 19 February 1985

The calcitonin (CT) gene is alternatively expressed in a tissue-specific fashion producing either the calcium regulatory hormone CT in the thyroid or the neuropeptide calcitonin gene related peptide (CGRP) in the brain. In medullary carcinoma of the thyroid both peptides are produced. We present here evidence for the existence in the human genome of a second CT gene, which is also expressed in human medullary thyroid carcinoma. This gene encodes a second human CGRP, differing from the known human CGRP in 3 of the 37 amino acids.

Human calcitonin *Calcitonin gene related peptide* *Medullary thyroid carcinoma*
Second human calcitonin gene *Second human calcitonin gene related peptide*

1. INTRODUCTION

Studies of the rat calcitonin (CT) gene have revealed that this gene generates two distinct mRNAs by the selective use of alternative polyadenylation sites within the primary transcript [1]. Thyroidal C-cells predominantly produce mRNA encoding the precursor of the calcium regulatory hormone calcitonin. This mRNA was shown to contain sequences transcribed from exons 1 to 4 of the rat CT gene. In the brain, spinal cord, and cranial nerve ganglia mRNA encoding the precursor of the neuropeptide calcitonin gene related peptide (CGRP) is the product of CT gene expression. Exons 1–3 of the rat CT gene are represented in both CT- and CGRP mRNA, whereas CGRP mRNA contains transcripts of exons 5 and 6 instead of exon 4 as its 3′-terminal part.

Structural analysis of the human calcitonin (hCT) gene (fig.1) suggests that this gene, like the rat gene, may be alternatively expressed in a tissue-specific fashion. The first 4 exons specify the mature hCT mRNA [2], and a fifth exon encoding hCGRP has been identified [3]. Using exon 5 sequences as probe, hCGRP mRNA has been detected in human medullary thyroid carcinoma (MTC), and precursor RNA species containing hCT- as well as hCGRP specific sequences were shown to be present in this tissue [3].

CGRP has been detected immunochemically in human thyroid, pituitary, and brain [4]. It has been isolated from human MTC tissue and characterized [5]. CGRP has been shown to have potent effects on the heart [6], and peripheral blood vessels as vasodilator [7].

To analyze the structure of hCGRP mRNA, and in particular the counterpart of rat exon 6 of the calcitonin gene, we screened a cDNA library of human MTC mRNA with a genomic hCGRP specific probe. The inserts of clones hybridizing to this probe were subjected to nucleotide sequence analysis.

In addition to hCGRP cDNA a clonal insert was detected which predicts the existence of a second CT/CGRP gene in the human genome.

2. MATERIALS AND METHODS

2.1. *cDNA library*

Total cellular RNA was isolated from a human MTC metastasis and poly(A)-rich RNA was ob-

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/85/$3.30

tained by oligo(dT)-cellulose chromatography. cDNA was synthesized using reverse transcriptase (Life Sciences) and oligo(dT)$_{12-18}$ as primer. ss-cDNA was converted to ds-cDNA using Klenow DNA polymerase. After S_1-nuclease treatment, about 15 dC residues per 3′-terminus were added using terminal transferase. The dC-tailed ds-cDNA was annealed to *Pst*I linearized pBR322 with 3′-terminal oligo(dG) tails. *E. coli* K12 1592 cells were transformed with the hybrid DNA and were grown on nitrocellulose filters on tetracycline containing agar plates. Replica filters were screened for colonies hybridizing to hCGRP specific probe labeled by nick translation.

2.2. *Nucleotide sequence analysis*

ds-DNA restriction enzyme fragments were labeled 5′-terminally using [γ-^{32}P]ATP and T_4-polynucleotide kinase, or 3′-terminally using [α-^{32}P]ddATP and terminal transferase. The fragments were digested with a second restriction endonuclease and were subsequently separated by 5% polyacrylamide gel electrophoresis. Nucleotide sequences were determined according to the chemical modification technique [8].

3. RESULTS

A cDNA library derived from human MTC metastasis mRNA was screened with an hCGRP (exon 5) specific probe (*Bgl*II-*Pvu*II fragment, fig.1). About one in 1000 cDNA clones hybridized to this probe. The inserts of two hybridizing clones, pCGRP-3 and pCGRP-4, with lengths of 1100 and 800 nucleotides, respectively, were subjected to nucleotide sequence analysis. Fig.2 shows the restriction enzyme recognition sites relevant to this analysis. Total plasmid DNA was digested with either *Pst*I, *Bgl*II, *Pvu*II, *Sac*I or *Bst*NI. The resulting fragments were 5′- or 3′-terminally labeled and digested with one of the remaining restriction endonucleases. Fragments labeled at one of their termini were isolated and their sequences analyzed.

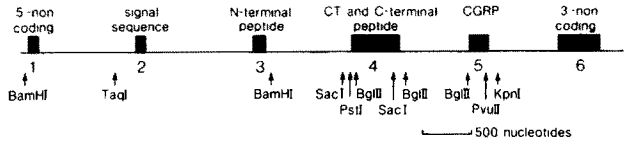


Fig.1. Organization of the hCT gene. The 6 exons and some restriction enzyme recognition sites are indicated. Gene organization was determined partially by DNA sequencing and by electron-microscopic heteroduplex mapping of a 6.5 kb *Taq*I fragment with hCT and hCGRP cDNAs, respectively [2].

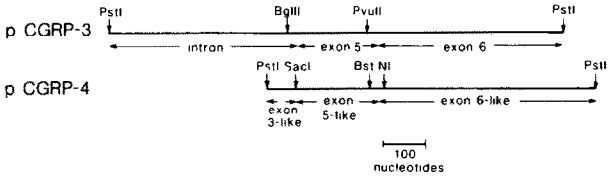
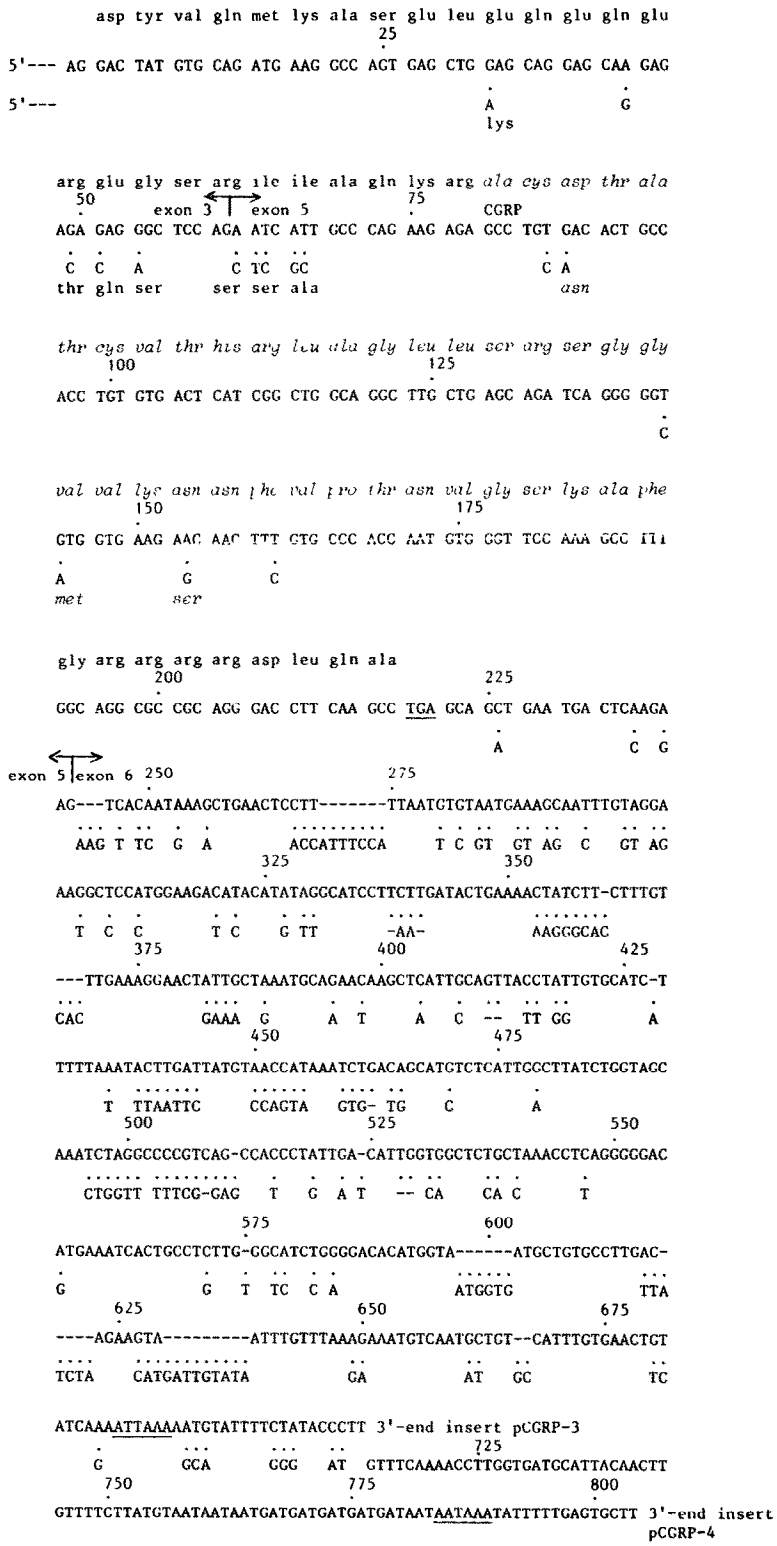


Fig.2. Restriction enzyme recognition sites within the inserts of plasmids pCGRP-3 and pCGRP-4. Nucleotide sequences were determined starting from these recognition sites after 3′- and 5′-end labeling of the fragments and digestion with a second restriction endonuclease. The horizontal arrows indicate from which parts of the hCT gene the sequences derive (pCGRP-3) or to which parts of the hCT gene the sequences are related (pCGRP-4).

The sequencing results show that the insert of pCGRP-3 is derived from the hCT gene. In addition to the sequence of exon 5 [3], a sequence of 432 nucleotides at the 3′-end was determined. This sequence is 80% homologous to the corresponding 3′-terminal non-translated exon 6 derived sequence in rat CGRP mRNA [1]. The sequence contains the variant polyadenylation enzyme recognition sequence ATTAAA (nucleotides 688–693 in fig.3), at the same position as in rat CGRP mRNA. On the basis of this homology, this sequence has been designated exon 6 in the hCT gene (fig.1). At

→

Fig.3. Comparison of the nucleotide sequence of the insert of pCGRP-4 to the sequence of exon 3 (derived from the hCT gene and hCT cDNA clones, fig.2), and exons 5 + 6 (derived from the hCT gene and pCGRP-3). Gaps (indicated by −) have been introduced to maximize nucleotide sequence homology. The amino acids encoded by the hCT gene and substitutions encoded by pCGRP-4 are shown. The hCGRP sequence is in italics. The TGA stop codon and the presumed poly(A) addition enzyme recognition sequences are underlined.


asp tyr val gln met lys ala ser glu leu glu gln glu gln glu
hCT gene 5'--- AG GAC TAT GTG CAG ATG AAG GCC AGT GAG CTG GAG CAG GAG CAA GAG
pCGRP-4 5'---
lys
arg glu gly ser arg ile ile ala gln lys arg ala cys asp thr ala
exon 3
exon 5
CGRP
AGA GAG GGC TCC AGA ATC ATT GCC CAG AAG AGA GCC TGT GAC ACT GCC
thr gln ser ser ser ala asn
thr cys val thr his arg leu ala gly leu leu ser arg ser gly gly
ACC TGT GTG ACT CAT CGG CTG GCA GGC TTG CTG AGC AGA TCA GGG GGT
val val lys asn asn phe val pro thr asn val gly ser lys ala phe
GTG GTG AAG AAC AAC TTT GTG CCC ACC AAT GTG GGT TCC AAA GCC TTT
met ser
gly arg arg arg arg asp leu gln ala
GGC AGG CGC CGC AGG GAC CTT CAA GCC TGA GCA GCT GAA TGA CTCAAGA
exon 5 exon 6
AG---TCACAATAAAGCTGAACTCCTT-------TTAATGTGTAATGAAAGCAATTTGTAGGA
AAGGCTCCATGGAAGACATACATATAGGCATCCTTCTTGATACTGAAAACTATCTT-CTTTGT
---TTGAAAGGAACTATTGCTAAATGCAGAACAAGCTCATTGCAGTTACCTATTGTGCATC-T
CAC GAAA G A T A C -- TT GG A
TTTTAAATACTTGATTATGTAACCATAAATCTGACAGCATGTCTCATTGCCTTATCTGCTAGC
T TTAATTC CCAGTA GTG- TG C A
AAATCTAGGCCCCGTCAG-CCACCCTATTGA-CATTGGTGGCTCTGCTAAACCTCAGGGGGAC
CTGGTT TTTCG-GAG T G A T -- CA CA C T
ATGAAATCACTGCCTCTTG-GGCATCTGGGGACACATGGTA------ATGCTGTGCCTTGAC-
G G T TC C A ATGGTG TTA
----AGAAGTA---------ATTTGTTTAAAGAAATGTCAATGCTGT--CATTTGTGAACTGT
TCTA CATGATTGTATA GA AT GC TC
ATCAAAATTAAAAATGTATTTTCTATACCCTT 3'-end insert pCGRP-3
G GCA GGG AT GTTTCAAAACCTTGGTGATGCATTACAACTT
GTTTTCTTATGTAATAATAATGATGATGATGATAATAATAAATATTTTTGAGTGCTT 3'-end insert pCGRP-4

the 5′-end of the insert of pCGRP-3, a sequence of 466 nucleotides corresponding to the region of the intron between exons 4 and 5 (the entire sequence of which has been determined, unpublished) adjacent to exon 5 was found. This probably indicates that the insert of pCGRP-3 is a product of reversed transcription of an incompletely processed hCGRP precursor RNA possessing a poly(A) tail. Such precursor molecules have been detected in large amounts in human MTC tissue [3]. It has been shown that in mature hCGRP mRNA exon 5 sequences are coupled to exon 3 sequences at their 5′-end, as they are in rat CGRP mRNA.

The insert of pCGRP-4 does not derive from the first hCT gene. Within this plasmid, a sequence 92% homologous to exon 3 of the hCT gene, followed by a sequence 92% homologous to exon 5 and a sequence 65% homologous to exon 6 were found (fig.3). At the 3′-end, a sequence of 91 nucleotides found in neither pCGRP-3 nor rat exon 6 is present. Although the variant ATTAAA sequence (nucleotides 688–693 in fig.3) is conserved, it is more likely that the usual AATAAA sequence at positions 784–789 serves as poly(A) addition enzyme recognition site within this mRNA.

4. DISCUSSION

The sequencing results indicate that the human genome contains, in addition to the known hCT gene, a second distinctly different hCT gene, which is expressed in human MTC. Evidence for the existence of a second hCT gene has already been obtained in hybridization studies of hCT gene specific probes with human DNA. In *Taq*I digested human DNA a 6.5 or 8.0 kb polymorphic restriction enzyme fragment reacts to a probe containing exons 2, 3, and 4 of the hCT gene [10]. In addition, 2.3 and 3.0 kb bands were observed which could not be explained as fragments of the hCT gene. A probe containing only exon 4 (hCT-encoding) sequences does not detect these extra bands. The exon 3-like part of pCGRP-4 (*Pst*I/*Sac*I fragment, fig.2) hybridizes to the 2.3 and 3.0 kb *Taq*I fragments in human DNA digests, and also to the 6.5/8.0 kb bands, as expected on the basis of its high degree of homology to the hCT gene exon 3. The exon 5 +6 part of pCGRP-3 hybridizes to the polymorphic 6.5/8.0 kb *Taq*I fragments, and, more weakly, to a 2.6 kb *Taq*I fragment; the exon 5 + 6-like part of pCGRP-4 (*Sac*I/*Pst*I fragment) hybridizes much stronger to the 2.6 kb *Taq*I fragment, and, more weakly, to the 6.5/8.0 kb *Taq*I fragments (fig.4). These 2.3, 2.6, and 3.0 *Taq*I fragments are not present within the inserts of both cosmid clones containing the hCT gene isolated by us [2]. This indicates that the second hCT gene is not located immediately adjacent to the hCT gene on chromosome 11 [10]. The 2.3, 2.6, and 3.0 kb *Taq*I fragments do not derive from an allele of the first hCT gene. In the DNA of all individuals screened for the 6.5/8.0 kb *Taq*I restriction enzyme recognition site polymorphism, the extra *Taq*I fragments are detected, in addition to fragments of 8.0 kb, 6.5 kb, or both [10].

It is noteworthy that, although this newly discovered gene shares a great deal of structural features with the hCT gene, a region homologous to the hCT encoding exon 4 seems to be absent. This is implicated by the absence of hybridizing *Taq*I fragments other than the 6.5/8.0 kb fragments to the exon 4 specific probe. Nevertheless, a second calcitonin has been reported to be produced in man [11,12]. This calcitonin-like peptide is indistinguishable from synthetic salmon calcitonin-(1–32) (sCT) on high-performance liquid chromatography. Since hCT and sCT are only 50% homologous in their amino acid sequences, although sCT has a 10-times more potent

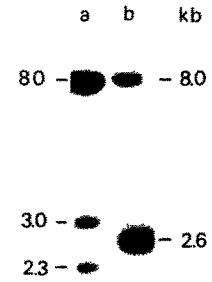


Fig.4. Southern blot analysis of human *Taq*I digested DNA. (a) A probe containing hCT gene exons 2, 3, and 4 detects an 8.0 kb fragment and hybridizes more weakly to 2.3 and 3.0 kb fragments. (b) A probe containing the exon 5+6-like sequences of the second hCT gene hybridizes strongly to a 2.6 kb fragment and to the 8.0 kb fragment, containing exons 2–6 of the first hCT gene.

hypocalcaemic effect in man than has hCT [13], it is possible that the coding sequence for the sCT-like substance in man is insufficiently homologous to the hCT coding sequence to be detected with a hCT specific probe. The possibility that the second hCT gene contains the code for the sCT-like material should be borne in mind.

The most intriguing aspect of this novel gene is the fact that it encodes a secnd hCGRP, differing from the known hCGRP in 3 of 37 amino acids (fig.3). Interestingly, 2 of these 3 differences occur at positions also modified in rat CGRP when compared to hCGRP, one of the modifications being the same (Asp vs Asn, residue 3). This probably indicates that these amino acid residues may be substituted by others without altering the biological activity of the peptide. The coding sequences for pairs of basic amino acids flanking the hCGRP sequence, needed for excision of hCGRP from its precursor, the Gly-Arg-Arg-Arg sequence needed for C-terminal amidation of hCGRP and the TGA stop codon, followed by a second TGA in the same reading-frame, are all conserved in the second hCT gene.

It is not unlikely that antibodies raised against hCGRP cross-react with the second hCGRP peptide specified by the second hCT gene. Immunochemical data demonstrating the presence of hCGRP in different tissues should therefore be interpreted with caution in respect to conclusions concerning the expression of the hCT gene.

REFERENCES

[1] Amara, S.G., Evans, R.M. and Rosenfeld, M.G. (1984) Mol. Cell Biol. 4, 2151–2160.
[2] Steenbergh, P.H., Höppener, J.W.M., Zandberg, J., Cremers, A.F.M., Jansz, H.S. and Lips, C.J.M. (1985) Proc. Int. Symp. Calcitonin 1984 (Pecile, A. ed.) pp.23–32, Elsevier, Amsterdam, New York.
[3] Steenbergh, P.H., Höppener, J.W.M., Zandberg, J., Van de Ven, W.J.M., Jansz, H.S. and Lips, C.J.M. (1984) J. Clin. Endocrinol. Metab. 59, 358–360.
[4] Tschopp, F.A., Tobler, P.H. and Fischer, J.A. (1984) Mol. Cell Endocrinol. 36, 53–57.
[5] Morris, H.R., Panico, M., Etienne, T., Tippins, J., Girgis, S.I. and MacIntyre, I. (1984) Nature 308, 746–748.
[6] Etienne, T. (1984) J. Physiol. 351, 48P.
[7] Brain, S.D., Williams, T.J., Tippins, J.R., Morris, H.R. and MacIntyre, I. (1985) Nature 313, 54–56.
[8] Maxam, A.M. and Gilbert, W. (1977) Proc. Natl. Acad. Sci. USA 74, 560–564.
[9] Nelkin, B.D., Rosenfeld, K.I., DeBustros, A., Leong, S.S., Roos, B.A. and Baylin, S.B. (1984) Biochem. Biophys. Res. Commun. 123, 648–655.
[10] Höppener, J.W.M., Steenbergh, P.H., Zandberg, J., Bakker, E., Pearson, P.L., Geurts van Kessel, A.H.M., Jansz, H.S. and Lips, C.J.M. (1984) Hum. Genet. 66, 309–312.
[11] Fischer, J.A., Tobler, P.H., Henke, H. and Tschopp, F.A. (1983) J. Clin. Endocrinol. Metab. 57, 1314–1316.
[12] Tobler, P.H., Tschopp, F.A., Dambacher, M.A. and Fischer, J.A. (1984) Clin. Endocrinol. 20, 253–259.
[13] Galante, L., Joplin, G.F., MacIntyre, I. and Woodhouse, N.J.Y. (1973) Clin. Sci. 44, 605–610.