Volume 226, number 1, 96-100 FEB 05415 December 1987

The structure of procalcitonin of the salmon as deduced from its cDNA sequence

E. Pöschl, I. Lindley, E. Hofer, J.-M. Seifert, W. Brunowsky and J. Besemer

Sandoz Forschungsinstitut, Brunner Straße 59, A-1235 Wien, Austria

Received 22 September 1987; revised version received 4 November 1987

Oligonucleotide probes based on the known amino acid sequence of salmon calcitonin were used to screen a cDNA library obtained from ultimobranchial glands of salmon for clones encoding salmon calcitonin. From the cDNA sequence of strongly hybridizing clones the complete primary structure of the calcitonin precursor could be deduced. Its overall structure is identical with the structures of procalcitonins from other vertebrates and has the highest homology with the chicken precursor.

Calcitonin; cDNA cloning; Nucleotide sequence; Amino acid sequence; (Salmon)

1. INTRODUCTION

The hormone calcitonin, a 32 amino acid peptide, is involved in the regulation of calcium metabolism and is secreted by C-cells which are of neuroectodermal origin and are found in the thyroid in mammals and in the ultimobranchial gland of fish, amphibians and birds. Calcitonin of higher and lower vertebrates is synthesized as a precursor from which it is cleaved by (a) proteolytic enzyme(s). In chicken, rat and human the calcitonin gene encodes a second hormone, a 37 amino acid neuropeptide, calcitonin gene related peptide (CGRP), which is also processed from a precursor. In rat and human synthesis of calcitonin and CGRP from one gene is regulated by a tissue specific alternate mRNA splicing event [1-3]. The very similar structures of mammalian and avian chromosomal genes for calcitonin and CGRP [4] suggest that regulation by alternate splicing occurs also in the chicken, although there is no evidence as yet for this event. The structure of salmon calcitonin has been known since 1969 [5] and the chemically synthesized hormone is in clinical use to treat Paget's disease and osteoporosis. The gene, however, which codes for calcitonin in the salmon is not yet characterized. Such characterization could reveal whether salmon utilizes alternate splicing to produce a CGRP, and help to elucidate the nature of the immunoreactive salmon calcitonin-like material detected in man [6,7], and produced in vitro from mRNA of human medullary carcinoma cells [4].

We report here the isolation and sequence of the cDNA coding for salmon calcitonin-I precursor.

Correspondence address: J. Besemer, Sandoz Forschungsinstitut, Brunner Straße 59, A-1235 Wien, Austria

The nucleotide sequence presented here has been submitted to the EMBL/GenBank database under the accession number Y00765

2. MATERIALS AND METHODS

2.1. *Tissues*

Ultimobranchial glands were collected from freshly caught salmon during the first two weeks of June 1984 at the Salmon farm of Froeya Fiskeindustri in Dyrvik, Norway. Immediately after

Published by Elsevier Science Publishers B.V. (Biomedical Division)

removal they were frozen and stored in liquid nitrogen.

2.2. *cDNA library construction, screening and DNA sequencing*

Total RNA was extracted from 70 g of ultimobranchial glands and surrounding connective tissue using guanodinium thiocyanate [8]. After oligo(dT)-Sepharose selection, the poly(A)$^+$ RNA was converted into cDNA using AMV reverse transcriptase, RNase H and DNA polymerase I [9,10]. Nuclease S_1 and T_4 DNA polymerase were used to trim the ends, which were oligo(dC) tailed and the cDNA was ligated into *Pst*I-cut oligo(dG)-tailed pBR322 and transformed into *E.coli* 5K cells. A total of 50 000 transformants was replica plated on nitrocellulose filters and screened by hybridization with two oligonucleotide probes. One, EH7, was a mixture of 48 14-mers, the degenerate sequence of which corresponded to amino acids 14–18 of calcitonin (fig.1), the other, EH8, was a 54-mer with a unique sequence in which degenerated codon positions were selected according to codon usages in other eucaryotes [11,12] and contained codons for amino acids 1–18 (fig.1). Oligonucleotides were labelled with T_4 polynucleotide kinase and [γ-^{32}P]ATP. After hybridization which followed standard procedures [9] filters were washed according to Wood et al. [13]. Plasmid DNA from colonies which gave positive signals with both probes EH7 and EH8 were purified by the boiling method [9] and were digested with restriction endonuclease *Pst*I. DNA fragments were separated on agarose gels and transferred to nylon filters. The fragments were hybridized to probe EH8 as described for colony hybridization.

For sequence analysis DNA was *Pst*I cut and subcloned into the *Pst*I site of bacteriophage M13mp10, and was sequenced by the chain termination method [14].

3. RESULTS

From the primary and secondary screening of the cDNA library five independently derived clones were identified, the DNA of which hybridized with both probe EH7 and EH8. Plasmid DNA from these clones was prepared and digested with *Pst*I, and the DNA fragments were separated by agarose gel electrophoresis. The insert lengths were found to be 4.2, 3.0, 2.0, 1.4 and 0.78 kb for the five clones (not shown). A detailed characterization of the five insert DNAs, including mapping of DNA with different restriction enzymes, hybridization of insert DNA to the EH8 probe and calcitonin cDNA and sequence analysis, revealed that none of these five cDNAs were related to each other and that only the clone with the DNA insert of 0.78 kb length (clone 404) contained calcitonin cDNA. The other clones contained sequences fully or partially homologous to oligonucleotide probes EH7 and EH8, but besides this showed no similarity to calcitonin cDNA. The clone with the insert of 3.0 kb contained a sequence of 28 bp which was only distantly related to the sequence of the EH8 probe, however, since it was repeated 18 times in tandem the cDNA gave strong hybridization signals with EH8 (not shown).

The 0.78 kb cDNA insert of clone no. 404 carried a *Pst*I site within its sequence, so that cutting with *Pst*I produced two fragments of 430 and 350 base pairs. These *Pst*I fragments were subcloned into bacteriophage M13mp10 and were sequenced. The DNA sequence and the deduced amino acid sequence are depicted in fig.1. When this cDNA sequence was analyzed it was apparent that it contained the complete procalcitonin sequence but not the information for the N-terminal methionine. The sequence of the clone 404 insert started with codons for Ala-Ala-Pro which in other calcitonin genes (fig.2) define amino acids at positions 24–26 of the calcitonin precursor. In clone 404 the 5′-portion of calcitonin cDNA is therefore absent. In order to find the missing part, the cDNA library was screened again by hybridization with radioactively labelled calcitonin RNA transcribed in vitro from the 430 bp *Pst*I fragment. A total of seven independently derived clones gave positive signals. Plasmid DNA was prepared from these clones, the inserts were cut out with *Pst*I and their lengths were determined by gel electrophoresis. Among these seven clones, three had *Pst*I fragments longer than 430 nucleotides, indicating that they could contain more sequence from the 5′-part of the cDNA. The longest insert of approx. 500 bp was sequenced. It carried the information for an in-frame methionine and an additional 23 amino acids leading into our already existing sequence. As

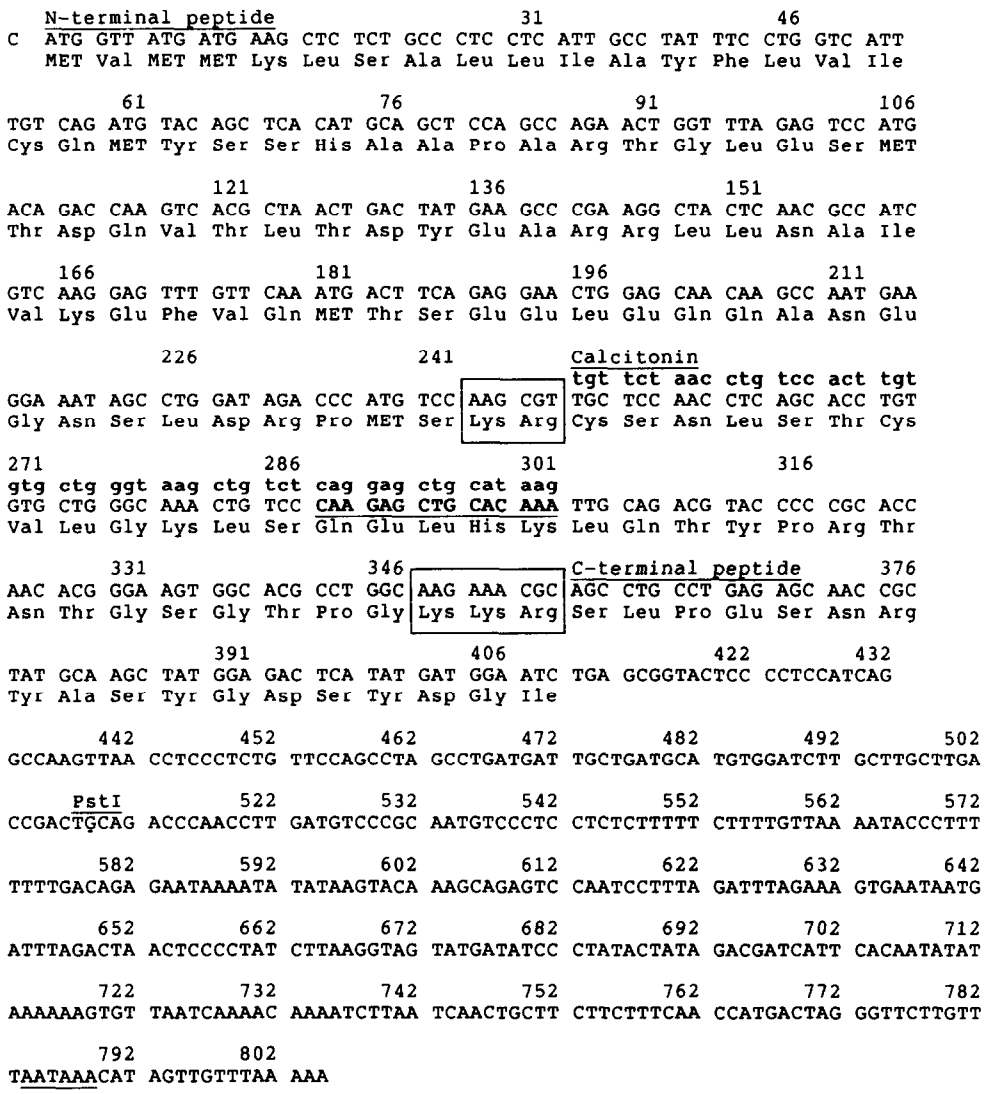

Fig.1. cDNA sequence of salmon procalcitonin and deduced amino acid sequence. The putative cleavage sites of the precursor are boxed. Hybridization probe EH8 is given in small letters, probe EH7 is underlined as is the putative adenylation signal at the 3'-end of the cDNA.

there was only one extra bp to the 5'-side of this methionine, it is possible that it is an internal amino acid. However, by comparison with the published structures of human and rat procalcitonin [1,20] it is extremely likely to be the starting methionine of the calcitonin precursor.

The amino acid sequence encoded by the cDNA was compared with the known procalcitonin sequences of human, rat and chicken [1–4]. From this comparison (fig.2) it is deduced that the precursor of salmon calcitonin has an N-terminal sequence of 80 amino acids followed by a putative cleavage site Lys-Arg which separates the N-terminal sequence from mature calcitonin. The calcitonin sequence is followed by a glycine which is necessary for amidation of the terminal proline of calcitonin and an additional cleavage site Lys-Lys-Arg. The C-terminal peptide is 18 amino acids long (fig.1). The cDNA characterized here encodes the salmon isohormone calcitonin-I [15].

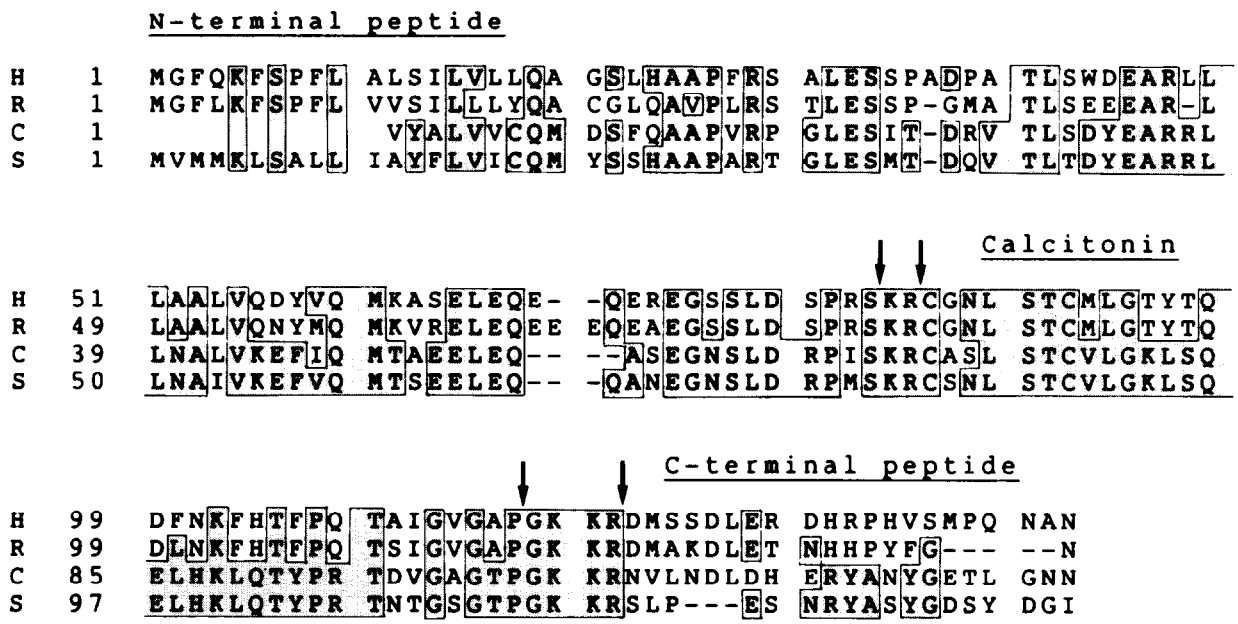

Fig.2. Comparison of amino acid sequences of calcitonin precursors of human (H), rat (R), chicken (C) and salmon (S). The 5'-terminal sequence of chicken precursor cDNA has not been published yet. Amino acids identical to the salmon precursor are boxed. Putative cleavage sites for the precursor are marked with asterisks. Sequences are from: H [20], Genbank Genetic Data Bank, accession J00109; R [1], EMBL/Database accession V01228; C [4], EMBL/Database accession X03012.

4. DISCUSSION

The deduced structure of salmon procalcitonin is very similar to those of the human, rat and chicken precursors. They all contain an N-terminal sequence of 78–82 amino acids, the calcitonin sequence of 32 amino acids and a C-terminal peptide of 16–21 amino acids. Comparison of these amino acid sequences (fig.2) shows that they are paired in terms of maximal homology: chicken and salmon show high homology, as do human and rat, but homology between the two pairs is lower. Gene evolution of the calcitonin precursors has been analyzed in more detail by Lasmoles et al. [4]. The authors came to the conclusion that the calcitonin sequences were strongly conserved during evolution up to the appearance of the mammals, and that the three different sequences comprising the precursor for calcitonin had had different rates of divergence during evolution, the calcitonin sequence being the most conserved. Our results are in agreement with these conclusions. Amino acid sequence homologies (fig.2) between salmon calcitonin and the three calcitonins of chicken, rat and human are 84% (27/32), 53% (17/32) and 50% (16/32), respectively. Similar numbers are obtained for the N-terminal peptide, namely 75% (52/69), 43% (34/80) and 50% (40/80), whereas the C-terminal peptide of the salmon has far less homology with the corresponding peptides of the other species. The functions of the N-terminal and C-terminal peptides of the procalcitonin are not known.

In the rat and human two different genes for two slightly different CGRPs, α and β, are expressed [16,17]. The β-CGRP gene of man also encodes in its exon 4 a calcitonin-like sequence [18,19]. However, this second calcitonin sequence seems not to be expressed in humans and may be a pseudogene. On the other hand it has been reported that a salmon-like calcitonin is expressed in humans [6,7] and mRNA directing the synthesis in vitro of a peptide which cross-reacts with antisera directed against salmon calcitonin was extracted from a medullary carcinoma [4]. With the salmon calcitonin cDNA in hand it will now be possible to

screen a cDNA library from medullary carcinoma and/or a human genomic library for the presence of salmon calcitonin-like DNA sequences.

ACKNOWLEDGEMENTS

We thank Dr Evelyn Lopez from the Laboratoire de Physiologie of the Museum National d'Histoire Naturelle in Paris for teaching us the art of ultimobranchial gland removal. We also thank Renate Hofer-Warbinek for subcloning of the calcitonin cDNA into the pSP64 vector DNA and F.J. Werner for assistance with software.

REFERENCES

[1] Amara, S.G., Jonas, V., Rosenfeld, M., Ong, E.S. and Evans, R.M. (1982) Nature 298, 240–244.
[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] Rosenfeld, M.G., Amara, S.G. and Evans, R.M. (1984) Science 225, 1315–1320.
[4] Lasmoles, F., Jullienne, A., Day, F., Minvielle, S., Milhaud, G. and Moukhtar, M.S. (1985) EMBO J. 4, 2603–2607.
[5] Niall, H.D., Keutmann, H.T., Copp, D.H. and Potts, J.T., jr (1969) Proc. Natl. Acad. Sci. USA 64, 771–778.
[6] Fischer, J.A., Tobler, P.H., Henke, H. and Tschopp, F.A. (1983) J. Clin. Endocrinol. Metab. 57, 1314–1316.
[7] Tobler, P.H., Tschopp, F.A., Dambacher, M.A. and Fischer, J.A. (1984) Clin. Endocrinol. 20, 253–259.
[8] Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J. and Rutter, W.J. (1979) Biochemistry 18, 5294–5299.
[9] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning, A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
[10] Gubler, U. and Hoffman, B.J. (1983) Gene 25, 253–269.
[11] Grautham, R., Gautier, C., Gouy, M. and Mercier, R. (1981) Nucleic Acids Res. 9, r43–r74.
[12] Lathe, R. (1985) J. Mol. Biol. 183, 1–12.
[13] Wood, W.I., Gitschier, J., Lasky, L.A. and Lawn, R.M. (1985) Proc. Natl. Acad. Sci. USA 82, 1585–1588.
[14] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[15] MacIntyre, I. and Stevenson, J.C. (1981) in: Calcitonin 1980 (Pecile, A. ed.) International Congress Series no. 540, p. 1–10, Excerpta Medica, Amsterdam, Oxford, Princeton.
[16] Amara, S., Arriza, J.L., Leff, S.E., Swanson, L.W., Evans, R.M. and Rosenfeld, M.G. (1985) Science 229, 1094–1097.
[17] Steenbergh, P.H., Höppener, J.W.M., Zandberg, J., Lips, C.J.M. and Jansz, H.S. (1985) FEBS Lett. 183, 403–407.
[18] Alevizaki, Shiraishi, A., Rassool, F.V., Ferrier, G.J.M., MacIntyre, I. and Legon, S. (1986) FEBS Lett. 206, 47–52.
[19] Steenberg, P.H., Höppener, J.W.M., Zandberg, J., Vissier, A., Lips, C.J.M. and Jansz, H.S. (1986) FEBS Lett. 209, 97–103.
[20] Craig, R.K., Hall, L., Edbrooke, M.R., Allison, J. and MacIntyre, I. (1982) Nature 295, 345–347.