Somatostatin-28 Encoded in a Cloned cDNA Obtained from a Rat Medullary Thyroid Carcinoma*

Richard H. Goodman‡§, John W. Jacobs‡§, Phillip C. Dee§, and Joel F. Habener‡§

From the §Laboratory of Molecular Endocrinology, Massachusetts General Hospital, Howard Hughes Medical Institute, Harvard Medical School, and the ‡Endocrine Division, Department of Medicine, New England Medical Center, Boston, Massachusetts 02114

We have constructed and cloned in bacteria complementary DNAs derived from a transplantable rat medullary thyroid carcinoma. Using a hybridization probe encoding an anglerfish islet pre-prosomatostatin, a precursor of the tetradecapeptide somatostatin, we have identified and isolated a clone containing a 400-base pair complementary DNA encoding most of the rat carcinoma pre-prosomatostatin. The amino acid sequence of the tetradecapeptide somatostatin and of the amino-terminally extended form, somatostatin-28 was deduced from the nucleotide sequence of the complementary DNA. Somatostatin-28 was found at the COOH terminus of a polypeptide of at least 80 amino acids indicating that somatostatin-28 arises by cleavage from a large precursor. The sequences of somatostatin-28 and somatostatin-14 are strictly conserved between the rat and other mammals. Such conservation of these sequences indicates strong selective pressures during evolution to maintain the sequence and suggests that somatostatin-28 may serve some essential biologic functions apart from, or in addition to, the important regulatory actions of somatostatin-14. Additionally, we found a high degree of homology in the amino acid sequences of the NH2-terminal extension peptides in the anglerfish islet and the rat carcinoma pre-prosomatostatins pointing further to a possible biologic function of these extension peptides.

Somatostatin is a tetradecapeptide that regulates the release of pituitary, pancreatic, and gastrointestinal hormones (1). Initially identified in the hypothalamus as an inhibitor of growth hormone secretion (2), somatostatin has subsequently been found in extrahypothalamic brain, spinal cord, retina, gastrointestinal tract, pancreatic islets, and thyroid (3–6). In addition to inhibiting the secretion of a number of peptide hormones, somatostatin has been proposed to act as a neurotransmitter and to modulate gastrointestinal motility (7, 8).

The diverse functions and the widespread distribution of the tetradecapeptide somatostatin (somatostatin-14) have focused attention on the biosynthesis of the hormone. Several studies have shown that somatostatin-14 is synthesized as part of a larger precursor. A 28-amino acid form of the hormone (somatostatin-28) has recently been identified in extracts of porcine hypothalamus (9), gastrointestinal tract (10), and ovine hypothalamus (11). Somatostatin-28 may have functions distinct from those of the tetradecapeptide (12-14).

* This work was supported in part by grants from the National Institutes of Health and the National Science Foundation. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Pulse and pulse-chase labeling studies in brain and pancreatic islets indicate that somatostatin-28 is also derived from a larger precursor (15, 16). Cell-free translations of mRNA isolated from the pancreatic islets and gastrointestinal tissues of anglerfish (17–19), the pancreatic islets of channel catfish (20) and rat hypothalamus (21) have confirmed the existence of large (Mr = 14,000 to 16,000) precursors of somatostatin (pre-prosomatostatins). Recently, nucleotide sequencing of cloned complementary DNAs (cDNAs) have provided the amino acid sequences of two anglerfish islet pre-prosomatostatins (22, 23). These sequences show that the somatostatin-14 peptides are located at the COOH termini of 119 to 121 amino acid precursors and that the fish and mammalian somatostatin-14 peptides have identical sequences. However, only partial conservation of the somatostatin-28 sequence was observed between fish and mammals. Little is known about the structures of the peptides that lie NH2-terminal to the somatostatin-28 sequence of mammalian pre-prosomatostatins.

We now report the use of a cloned cDNA containing coding sequences for an anglerfish islet pre-prosomatostatin as a hybridization probe to identify a cloned somatostatin-related cDNA derived from a transplantable rat medullary carcinoma of the thyroid. Nucleotide sequencing of the cDNA revealed the complete sequence of rat somatostatin-28 and the partial sequence of the NH2-terminal precursor extension of the pre-prosomatostatin. We find that the amino acid sequence of somatostatin-14 and somatostatin-28 in the rat are identical with those found in the ovine and porcine species. In addition, comparison of the amino acid sequences of the NH2-terminal extensions encoded in the 5′ regions of the fish and rat cDNAs show considerable homology (53% among nucleotides and 39% among amino acids). The latter observations point to the existence of strong selective pressures to conserve the structures of the NH2-terminal extensions over the 400 million years since fish and rat diverged in evolution. These findings suggest that these peptide extensions may have some, as yet unrecognized, biologic function other than simply to serve as protein "spacer" sequences.

EXPERIMENTAL PROCEDURES

Construction and Cloning of cDNAs from a Rat Medullary Carcinoma of the Thyroid—The preparation of a cloned cDNA library from a rat medullary carcinoma of the thyroid kindly provided by N. H. Bell was described previously (24). In brief, cDNA was prepared from the polyadenylated RNA (25) by using an oligo(dT) primer and reverse transcriptase (26). Double-stranded DNA was prepared from the cDNA with polymerase I, inserted into the Pst I restriction endonuclease site of the plasmid pBR322 (26), and recombinant plasmids were introduced into Escherichia coli χ1776 by the procedure of Villa Komaroff et al. (27).
of the procedure of Grunstein and Hogness (28). Cloned cDNA encoding anglerfish pre-prosomatostatin (22) was digested with the restriction endonuclease Xma I and Dde I, labeled at the 3'-end with 32P (22), and electrophoresed on 5% polyacrylamide gels containing Tris/borate/Na,EDTA. A 58-base pair DNA fragment encoding the tetradecapeptide somatostatin was isolated from the gels and used as a hybridization probe.

Bacterial colonies containing recombinant cDNAs derived from rat medullary thyroid carcinoma were grown on nitrocellulose filters. Bacteria were lysed and the DNA was fixed onto the filters as described by Grunstein and Hogness (28). Filters were prehybridized for 18 h at 68°C in a solution containing 6 × SSC (0.9 M NaCl, 0.09 M sodium citrate), 5 × Denhardt’s reagent (0.1% w/v each of bovine serum albumin, polyvinylpyrrolidone, and Ficoll), 0.5% sodium dodecyl sulfate, 5 μg/ml of sonicated denatured salmon sperm and E. coli DNA, and 10% dextran sulfate.

After prehybridization, the buffer was discarded and replaced by an identical solution containing heat-denatured hybridization probe. Filters were then washed to hybridize for 2 h at 68°C and were subsequently washed 10 times at 68°C with a solution containing 5 × SSC, 1 × Denhardt’s reagent, and 0.5% sodium dodecyl sulfate. Filters were then washed twice at room temperature in 0.01 × SSC and air-dried. Autoradiograms were prepared using Kodak X-O-Mat film exposed for 2 to 3 days at −90°C using a Dupont Cronex intensifying screen.

RESULTS

Immunoreactive somatostatin has been identified in extracts of transplantable rat medullary thyroid carcinomas (30, 31). Preliminary studies using immunoprecipitation and hybridization selection techniques were unsuccessful, however, in identifying somatostatin precursors from the cell-free translation products programed by medullary thyroid carcinoma mRNAs.1 We had previously constructed a cDNA library from a rat medullary thyroid carcinoma for the purpose of isolating and sequencing cDNA encoding a precursor of calcitonin (24). By using a cloned cDNA encoding a somatostatin precursor from anglerfish islets we were able to identify a rat somatostatin-related cDNA in the carcinoma cDNA library by the colony hybridization method. Only one clone, 400 base pairs in length, was identified out of approximately 2,000 clones which were screened. The paucity of clones containing somatostatin-related cDNAs reflects the low levels of soma-
tostatin found in these tumors (30).

The identification of sites which are cleaved by the restriction endonucleases Xma I, Pst I, and Rsa I proved particularly useful in determining the nucleotide sequence of the cDNA. Restriction fragments were sequenced on both the sense and nonsense strands (Fig. 1).

The nucleotide sequence of the cloned cDNA and the corresponding amino acid sequence is shown in Fig. 2. This cDNA encodes the sequence of the tetradecapeptide somato-
statin and somatostatin-28 in addition to a 51-amino acid NH2-terminal extension. As predicted by Patzl et al. (32), the tetradecapeptide somatostatin is located at the COOH terminus of the precursor, followed by a stop codon TAG. The 14 amino acids predicted by the nucleotide sequence which precede the tetradecapeptide are identical with those found in ovine and porcine somatostatin-28 (9-11).

Somewhat surprising was the nucleotide sequence coding for Leu-Gln-Arg which separates somatostatin-28 from the remainder of the precursor (Figs. 2 and 3). This sequence is reminiscent of the sequence Leu-Glu-Arg found at the analogous position in the anglerfish pre-prosomatostatins (23). In-

1 R. H. Goodman, unpublished data.
asmuch as the sites of cleavages typically found in prohor-
mones consist of combinations of two of three of the basic
amines acids arginine and lysine (33), our sequence raises
the question of whether somatostatin-28 is produced by post-
translational cleavages of prosomatostatin in the rat medul-
lothyroid carcinoma.

The sequence Asn-Gln-Thr within the NH2-terminal exten-
sion of the rat somatostatin precursors represents a potential
N-glycosylation site (34). Patzelt et al. have suggested that
rat prosomatostatin may be glycosylated (32). No analogous
potential glycosylation sites are present within the two an-
glerfish pre-prosomatostatins.

Inasmuch as there were differences between the sequences
of anglerfish pre-prosomatostatin I reported by our laboratory
(22) and that of Hobart and co-workers (23), we have exten-
sively reanalyzed the sequence of the anglerfish somatostatin
precursor. The revised sequence, which is slightly different
from either of the two sequences reported previously is shown
in Fig. 4.

Fig. 4 additionally compares the nucleotide and amino acid
sequences of the anglerfish and rat somatostatin precursors.
Within the coding sequence for the tetradecapeptide soma-
tostatin, 36 of the 42 nucleotides (83%) are conserved between
rat and angelfish. Within the coding sequence for somatosta-
tin-28, 22 of the 28 amino acids (79%), and 58 of the 84
nucleotides (69%) are maintained. The six amino acid substi-
tutions between the fish and rat somatostatin-28 sequences
appear to be conservative in nature. It is necessary to add or
delete specific codons to maintain the homology between the
NH2-terminal extensions of the fish and rat sequences. If these
deductions and additions are made, 20 of the 51 amino acids
within the extension (39%) are strictly conserved and an
additional 15 amino acid changes are conservative in nature.
Eighty-one of 153 nucleotides (53%) within this region are
conserved. With the exception of the AATAAA sequence and
the polyadenylate tail, there appears to be little conservation
of nucleotide sequence within the 3' untranslated regions.

DISCUSSION

A cloned cDNA of 400 base pairs coding for pre-prosoma-
tostatin, a precursor of rat somatostatin-28, was identified by
colonial hybridization using a 32P-labeled cDNA encoding an
anglerfish idlet pre-prosomatostatin. Previous studies have
indicated that rat hypothalamic and pancreatic somatostatin-14
are identical in amino acid composition and chromatographic
properties with those isolated from ovine and porcine hypothalamus and pigeon and angelfish pancreas (35, 36).

The nucleotide sequence of our cDNA indicates that the primary structure of somatostatin-14 in rat is identical with
that of these other species. This sequence differs from that of
catfish islet somatostatin (37) and angelfish islet somatostatin
II (23). The amino acid sequence derived from the nucleotide
sequence of our cDNA indicates that the somatostatin-like
immunoreactive material isolated from the medullary thyroid
carcinoma is authentic somatostatin.

The cDNA which we have characterized encodes the entire
3' untranslated region, the sequence of somatostatin-28 at the
COOH-terminal of the precursor, and a portion of the NH2-
terminal peptide extension. Estimates of the size of prosoma-
tostatin isolated from rat pancreatic islets (32) and of pre-
prosomatostatin from cell-free translations of rat hypotha-
lamic mRNA (21), indicate that rat pre-prosomatostatin con-
tains approximately 110 amino acids. Inasmuch as the NH2-
terminal leader or signal sequence of the pre-prohormone
probably includes 24 to 28 amino acids, our partial cDNA
encoding 79 amino acids represents nearly the entire rat
prosomatostatin sequence.

Evidence regarding the processing of prosomatostatin in
the rat medullary thyroid carcinoma has been conflicting. Berelwitz et al. (30) have suggested that the predominant
form of somatostatin in their tumor was the tetradecapeptide.
Benoit et al. (38) concluded that somatostatin-28 constituted
the major immunoreactive form of the hormone. It was of
interest therefore to examine the peptide sequence adjacent
to somatostatin-28. The sequence Gln-Arg predicted from our
cDNA differs from that of a typical prohormone cleavage site.
It is possible that such a change in amino acid sequence at
this site might affect the relative efficiency of processing to
somatostatin-28 or somatostatin-14. It is conceivable in fact
that the regulation of the processing of multigenic hormones
such as somatostatin may involve production of separate
prohormones with different cleavage sites. Further studies
correlating the sequences of somatostatin cDNAs derived
from particular tissues with the somatostatin-28 to somatosta-
tin-14 ratios should elucidate this hypothesis. The possibility
of a reverse transcriptase error during the construction of the
cDNA library must also be considered. Although we were
unable to confirm the nucleotide sequence on an independ-
ently cloned cDNA due to the low abundance of the pre-
prosomatostatin cDNAs in the cDNA library prepared from
the carcinoma, eventual determination of the genomic se-
quence should provide such confirmation.

The revised sequence of the anglerfish pre-prosomatostatin
cDNA depicted in Fig. 4 more closely resembles that of Hobart
et al. (23) than the sequence which we initially reported (22).
This difference is primarily due to a sequencing error resulting
in a frame shift involving 17 of the 121 codons. Nonetheless,
there remain a few codons which differ from those reported
by Hobart et al. (23). These differences could result from
reverse transcription errors or from the sequencing of distinct
polymorphic cDNAs.

The sequence of porcine and ovine somatostatin-28 has
previously been determined (9-11). Recent evidence suggests that the biologic activities of somatostatin-28 may be greater than, and perhaps different from, those of somatostatin-14 (12-14). Strict conservation of the amino acid sequence of somatostatin-28 between the rat and these other mammals, species which diverged over 75 million years ago (39), is strong evidence for the existence of evolutionary pressures to maintain this sequence. It is likely, therefore, that somatostatin-28 has an important biologic function in mammals. Conservation between anglerfish and mammals of the six amino acids adjacent to the tetradecapeptide somatostatin suggests a particular importance of this portion of somatostatin-28 (Fig. 5).

Examination of the NH2-terminal portions of the fish and rat somatostatin precursors reveals several additional regions of considerable homology. This observation raises the possibility that the NH2-terminal portion of prosomatostatin may also have some specific biologic functions. Comparison of somatostatin cDNAs from other species should be useful in understanding the importance of this region. We are currently synthesizing by chemical methods peptide fragments of the NH2-terminal region of rat prosomatostatin for the preparation of antiserum to be used in studies of intracellular transport, secretion, and potential biologic activity of the precursor region of the prosomatostatin. Such studies should further our understanding of neuropeptide biosynthesis and physiology.

Acknowledgments—We thank N. H. Bell for providing the frozen rat carcinoma used initially in the preparation of RNA, W. C. Chin, P. K. Lund, and G. Heinrich for helpful advice, and J. Sullivan for help in the preparation of the manuscript.

REFERENCES

Somatostatin-28 encoded in a cloned cDNA obtained from a rat medullary thyroid carcinoma.

R H Goodman, J W Jacobs, P C Dee and J F Habener