Proc. Natl. Acad. Sci. USA
Vol. 87, pp. 4330–4334, June 1990
Biochemistry

cDNA cloning of somatolactin, a pituitary protein related to growth hormone and prolactin

(teleost fish/N-terminal sequence/gene family/evolution)

MASAO ONO*†, YOSHINAGA TAKAYAMA*, MARIANN RAND-WEAVER‡, SHUSAKU SAKATA§, TERUO YASUNAGA¶, TOYOHIKO NOSO‡, AND HIROSHI KAWAUCHI‡

*Department of Molecular Biology, School of Medicine, Kitasato University, Sagamihara, Kanagawa 228, Japan; ‡Laboratory of Molecular Endocrinology, School of Fisheries Sciences, Kitasato University, Sanriku, Iwate 022-01, Japan; §Central Technical Research Laboratory, Nippon Oil Co., Ltd., Naka-ku, Yokohama, Kanagawa 231, Japan; and ¶Computation Center, Institute of Physical and Chemical Research, Wako, Saitama 351, Japan

Communicated by Howard A. Bern, March 19, 1990 (received for review January 5, 1990)

ABSTRACT **From a flounder pituitary cDNA library, cDNA clones encoding a 28-kDa glycoprotein produced by the pars intermedia of the pituitary were isolated and characterized. Nucleotide sequencing demonstrated a precursor of the 28-kDa protein, which consisted of 231 amino acid residues, to be cleaved into a signal peptide (24 amino acids) and a mature protein (207 amino acids) containing one N-glycosylation site. By comparison of amino acid sequences, the 28-kDa protein was found to be distantly and similarly related to growth hormone and prolactin. Consequently, it was named somatolactin. Somatolactin mRNAs were specifically expressed as 1.2 and 1.8 kb poly(A)$^+$ RNAs in flounder pituitary.**

The pars intermedia of the teleost hypophysis contains two cell types that can usually be distinguished by staining with periodic acid/Schiff reagent (PAS) and lead/hematoxylin. Lead/hematoxylin-positive cells are the source of proopiomelanocortin-related peptides such as melanocyte-stimulating hormone (1), whereas the product of PAS-positive cells has not been characterized (2–5).

Histological studies have shown that the pars intermedia PAS-positive (PIPAS) cells can be activated under conditions of black background (2), acidic pH (3), low calcium (4), or low osmolarity (5) of the ambient water. The PIPAS cell response to these changes differed depending on the species but suggested that these cells may produce a hormone essential for adaptation to environmental changes. According to a biosynthetic study by van Eys *et al.* (6), these cells may produce proteins of 25 and 27 kDa, but further characterization has yet to be conducted.

Recently, a 26-kDa pituitary protein was isolated from the teleost Atlantic cod (*Gadus morhua*) and shown to be produced by PIPAS cells. Amino acid sequencing revealed it to be structurally related to growth hormone (GH) and prolactin (PRL) from the pituitary pars distalis (M.R.-W., T.N., B. I. Baker, and H.K., unpublished data). By use of the antibody raised against the cod 26-kDa protein, the corresponding 28-kDa protein was identified in flounder (*Paralichthys olivaceus*) pituitaries. Here we describe the cDNA cloning and structural analysis of the flounder 28-kDa protein.‖ Structural analysis indicated this protein to be distantly and similarly related to, but clearly distinct from, GH and PRL. These hormones along with placental lactogen (PL) comprise the GH/PRL family and may have been derived from a common ancestor by gene duplication events (7). The 28-kDa protein can be considered a new member of the GH/PRL family and has been tentatively designated as somatolactin in consideration of its structural resemblance to GH (somatotropin) and PRL.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

MATERIALS AND METHODS

Purification of Flounder 28-kDa Protein. Flounder (*P. olivaceus*) pituitary glands were homogenized and extracted under alkaline conditions (pH 9.5). The extract was chromatographed on Sephadex G-100 and the 28-kDa protein was purified by reverse-phase HPLC on TSK-gel ODS-120T column dimensions, 0.46 × 25 cm; particle size, 5 μm (unpublished data). The yield of flounder 28-kDa protein was 2.5 mg/g of pituitary (wet weight). N-terminal sequencing was performed using a Shimadzu PSQ1 protein sequencer.

Nucleic Acid Preparation and Filter Hybridization. Restriction and modification enzymes were obtained from Takara Shuzo (Kyoto) and New England Biolabs. Pituitaries and other tissues were obtained from flounders weighing about 0.5 kg. The preparation of RNA, isolation of poly(A)$^+$ RNA, and plaque hybridization were carried out as described (8). DNA was prepared from a single female flounder and chum salmon liver (8). Filter hybridization was conducted as described (9). Oligodeoxynucleotide probes were synthesized by an Applied Biosystems model 381A synthesizer and each was labeled at the 5′ end with T4 polynucleotide kinase and [γ-^{32}P]ATP.

cDNA Library, Screening, and Sequencing. A flounder pituitary cDNA library was constructed by the cDNA "synthesis system plus" and cDNA cloning system λgt10 (Amersham). From 0.2 μg of cDNA and 0.6 μg of λgt10 arm DNA, a library giving 5.1 × 10^4 plaques with 60% recombinants was obtained. Following plaque hybridization with ^{32}P-labeled oligonucleotide probes, the filter was washed four times (10 min per washing) in 1× SSC/0.1% SDS at 60°C (1× SSC is 0.15 M NaCl/0.015 M trisodium citrate, pH 7.6). The nucleotide sequences of cDNA were determined by the dideoxy chain-termination method (10), and all sequences in both DNA strands were determined. Labeling of cloned DNA was performed by the random-primed method (11) using [α-^{32}P]dCTP. The sequences were examined for similarity (12) to sequences in the GenBank and National Biomedical Research Foundation data bases (Releases 60.0 and 20.0, respectively) by using a Vax-11/780 (Digital Equipment) computer system at the Institute of Medical Science, University of Tokyo. Alignment of amino acid sequences was carried out following the introduction of gaps to maximize identity. For assessment of identity (I), each gap was counted as one substitution regardless of its length, and divergence (KC) was calculated as $KC = -\ln I$. An evolutionary tree was made from corrected divergence by the unweighted pair-group

Abbreviations: GH, growth hormone; PAS, periodic acid/Schiff reagent; PIPAS cells, pars intermedia PAS-positive cells; PL, placental lactogen; PRL, prolactin.

†To whom reprint requests should be addressed.

‖The sequences reported in this paper have been deposited in the GenBank data base (accession nos. M33695 and M33696).

clustering method (13), in which divergence between two proteins is expressed as height based on branching points.

RESULTS

Characterization of cDNA Clones. Based on the N-terminal amino acid sequence of the 28-kDa glycoprotein, 53-mer antisense oligonucleotide probes (nucleotide positions 103–155 in Fig. 1 *Lower*) were designed with human codon-usage data (14). Screening of the flounder pituitary cDNA library with these probes provided positive clones at a frequency of ≈1%. The 29 positive clones examined were grouped into 9 clones each with about 1 kb of cDNA and 20 clones each possessing about 1.6 kb of cDNA. Nucleotide sequences about 0.2 kb long were determined from both the 5′ and 3′ ends of 4 clones belonging to each group. Although the positions of the 5′ ends and the lengths of the poly(A) tails differed for each clone, overlapping sequences from both ends were identical in each group but sequences upstream from the poly(A) tail in the two groups were unrelated. Thus, the two clones cfSL3 and cfSL7 were selected. These possessed the longest cDNAs in each group and their nucleotide sequences were determined (Fig. 1 *Lower*).

The sequence of cfSL3 was identical to that of cfSL7 from the 5′ end of cfSL7 (position 8) to the polyadenylylation site (position 1031) of cfSL3. cfSL7 had an additional 577 bases with 15 consecutive adenine residues at the 3′ end. In both clones, sequences corresponding to the polyadenylylation signal AATAAA were found situated about 20 bases upstream from the polyadenylylation site.

The longest open reading frame (nucleotides 19–711), consisting of 231 codons, was present in both cfSL3 and cfSL7. Since the amino acid sequence specified by codons 25–47 of this open reading frame was identical with the N-terminal 23 residues of the 28-kDa protein (Fig. 1 *Lower*), both cfSL3 and cfSL7 were concluded to be cDNA clones for the 28-kDa protein. The N-terminal 24 residues encoded by the open reading frame formed a signal peptide, as evident from the abundance of hydrophobic amino acids (15). Consequently, the mature form of the 28-kDa protein was concluded to possess 207 amino acid residues and to have a molecular mass of 23,996 daltons. Glycosylation of the 24-kDa protein at a potential N-glycosylation site (Asn-Lys-Thr, residues 145–147) would increase the molecular mass; the apparent value for this parameter was found by SDS/PAGE to be 28 kDa. Downstream from nucleotide 712, no open

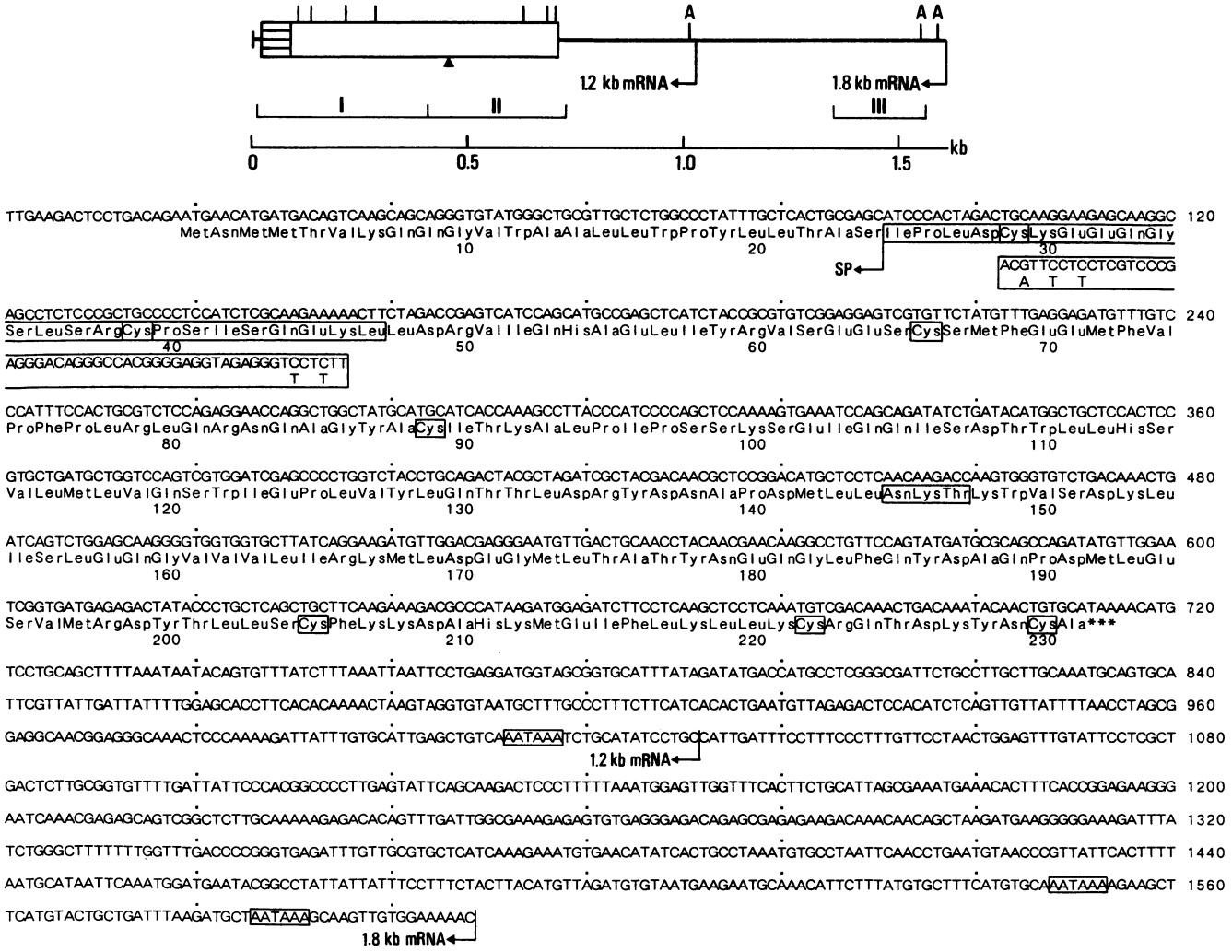

FIG. 1. Structure, nucleotide sequence, deduced amino acid sequence of cDNAs for flounder 28-kDa protein. (*Upper*) Bar represents the protein coding region. The position of the signal peptide is shown by horizontal striping. Vertical line, cysteine residue; ▲, potential N-glycosylation site; A, polyadenylylation signal; kb, kilobases. (*Lower*) Sequence of the mRNA-like DNA strand and deduced amino acid sequence of the precursor for the 28-kDa protein are shown. Amino acid residues are numbered from the putative start codon. The 53-mer oligonucleotide probes, N-terminal 23 residues of 28-kDa protein determined by protein sequencing, cysteine residues, potential N-glycosylation site, and polyadenylylation signals are boxed. SP, signal peptide. cfSL3 cDNA comprised nucleotides 1–1031 plus A_4. cfSL7 cDNA comprised nucleotides 8–1608 plus A_{15}.

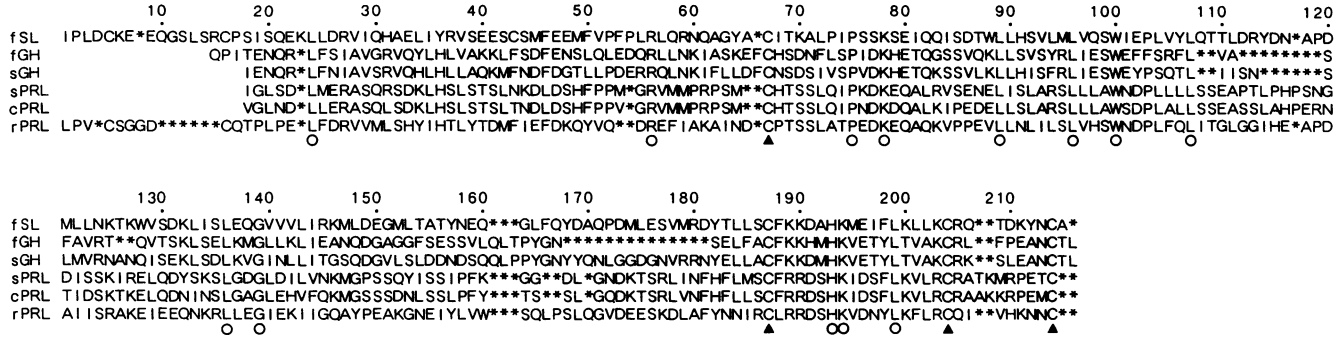

FIG. 2. Alignment of amino acid sequences of flounder 28-kDa protein (flounder somatolactin, fSL) with GH and PRL. Amino acids are represented by one-letter symbols. Gaps (asterisks) were inserted to enhance similarity. ▲, Conserved cysteine residue; ○, conserved amino acid. Reported amino acid sequences of flounder GH (fGH; ref. 16), chum salmon GH (sGH; ref. 17), chum salmon PRL (sPRL; ref. 18), carp PRL (cPRL; ref. 19), and rat PRL (rPRL; ref. 7) were used.

Table 1. Amino acid sequence identity and divergence of somatolactin and related proteins

	fSL	fGH	sGH	rGH	hGH	hPL	sPRL	cPRL	rPRL	hPRL	rPL
fSL		1.42	1.38	1.27	1.36	1.42	1.48	1.41	1.34	1.35	1.55
fGH	0.24		0.53	1.05	1.22	1.22	1.56	1.62	1.55	1.58	1.58
sGH	0.25	0.59		0.96	1.04	1.09	1.64	1.59	1.45	1.47	1.59
rGH	0.28	0.35	0.38		0.43	0.53	1.47	1.57	1.44	1.46	1.65
hGH	0.26	0.30	0.35	0.65		0.18	1.48	1.60	1.52	1.54	1.54
hPL	0.24	0.30	0.34	0.60	0.85		1.63	1.66	1.47	1.52	1.57
sPRL	0.23	0.21	0.19	0.23	0.23	0.20		0.31	1.20	1.06	1.34
cPRL	0.24	0.20	0.20	0.21	0.20	0.19	0.73		1.17	1.06	1.30
rPRL	0.26	0.21	0.24	0.24	0.22	0.23	0.30	0.31		0.47	1.04
hPRL	0.26	0.21	0.23	0.23	0.21	0.22	0.35	0.35	0.63		0.95
rPL	0.21	0.21	0.20	0.20	0.21	0.21	0.26	0.27	0.35	0.39	

Based on the aligned amino acid sequences, identity (*I*) and divergence (*KC*) were determined and are shown in the lower left and upper right halves of the table, respectively. Amino acid sequences of rat GH (rGH; ref. 7), human GH (hGH; ref. 7), human PL (hPL; ref. 20), human PRL (hPRL; ref. 7), and rat PL (rPL; ref. 21) were used. References and abbreviations for the other sequences are shown in the legend to Fig. 2.

reading frame capable of encoding more than 48 amino acids was found.

The 28-kDa Pituitary Protein Is Related to GH and PRL. A computer search for related sequences found that the 28-kDa protein bears significant structural similarity to GH and PRL (Fig. 2 and Table 1), thus indicating it to be a member of the GH/PRL family. However, the 28-kDa protein is distinct from GH since the extent of identity to flounder GH is only 24%. Compared with the close similarities between salmon and carp PRLs (73% identity) and between salmon and flounder GHs (59% identity), the 28-kDa protein is much less similar to these PRLs and GHs (average 24% identity) and thus distinct from PRL, although the amino acid sequence of flounder PRL has yet to be determined. Thus, the name somatolactin may be tentatively proposed for the 28-kDa protein in consideration of its distant relation to both GH (somatotropin) and PRL.

Somatolactin was found to have seven cysteine residues, four of which are in the C-terminal portion of the protein and are conserved in the GH/PRL family (7). The amino acid sequence of the N-terminal region of somatolactin and the location of the two cysteine residues in this region bear some resemblance to tetrapod PRL but not to fish PRL, which lacks the characteristic N-terminal disulfide loop (7). Phylogenetic relationships of the GH/PRL family including somatolactin and PLs were constructed on the basis of sequence comparisons (Table 1 and Fig. 3). Somatolactin was shown not to be closely related to PLs (24% and 21% identity to human and rat PLs, respectively) comprised in the GH/PRL family. Somatolactin is also distantly related to mouse proliferin (23% identity; ref. 22), a placental protein distantly related to GH and PRL. From the foregoing, we conclude that somatolactin is a pituitary protein belonging to the GH/PRL family but clearly distinct from GH, PRL, PL, and proliferin.

Expression and Organization of the Somatolactin Gene. The expression of the somatolactin gene in flounder pituitary was examined by Northern hybridization (Fig. 4*A*). Both probe I and probe II, mainly encoding the N-terminal half and C-terminal half of somatolactin, respectively, hybridized with 1.2-, 1.8-, 3.2-, and 3.8-kb RNAs, whereas probe III, encoding the 3′ untranslated region specific for cfSL7, hybridized only with 1.8- and 3.8-kb RNAs. It thus follows that clone cfSL3 derives from 1.2-kb somatolactin mRNA whereas cfSL7 derives from 1.8-kb somatolactin mRNA. The 3.2- and 3.8-kb RNAs are likely to be intron-containing precursor

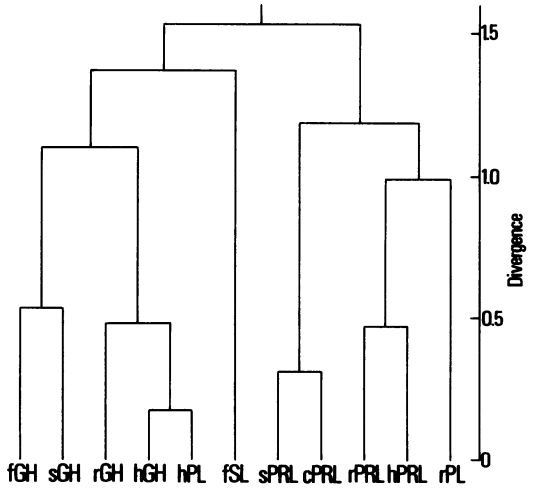

FIG. 3. Evolutionary tree of somatolactin and related proteins. The evolutionary tree was made from corrected divergence in Table 1 by the unweighted pair-group clustering method (13). Abbreviations for hormones are shown in the legend to Fig. 2 and Table 1.

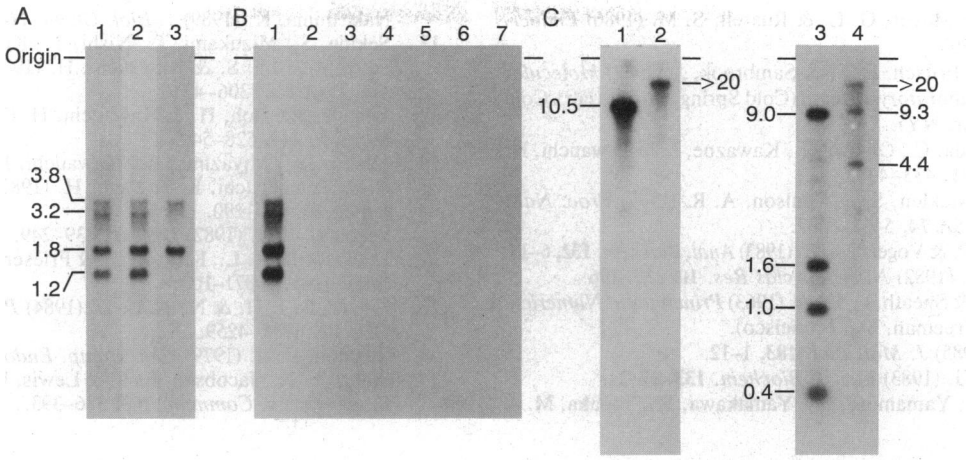

FIG. 4. Expression and organization of the somatolactin gene. (*A* and *B*) Expression of somatolactin mRNA. Total RNA (3 μg per lane) was denatured, fractionated in a 1% agarose gel containing formaldehyde, transferred to a filter, and hybridized with probe I (nucleotides 8–405 in Fig. 1 *Lower*) (*A* lane 1 and all lanes in *B*), probe II (nucleotides 406–725) (*A* lane 2), or probe III (nucleotides 1346–1558) (*A* lane 3). Following hybridization at 37°C, the filter was washed in 0.3× SSC/0.1% SDS at 65°C. Intactness of total RNA was confirmed by agarose gel electrophoresis following denaturation with glyoxal. RNA samples were from flounder pituitary (all lanes in *A* and lane 1 in *B*) and from flounder stomach, liver, heart, intestine, spleen, and brain (*B* lanes 2–7, respectively). (*C*) Organization of the somatolactin gene. *Eco*RI-digested (lanes 1 and 2) or *Pst* I-digested (lanes 3 and 4) liver DNA (2 μg) was analyzed by Southern hybridization with probes I and II. Following hybridization, the filter was washed in 1× SSC/0.1% SDS at 65°C. Lanes: 1 and 3, flounder; 2 and 4, chum salmon. Fragment sizes, indicated in kilobases, were estimated by a comparison with *Hin*dIII digests of phage λ DNA.

RNAs for 1.2- and 1.8-kb somatolactin mRNAs, respectively. The above findings confirm the alternative polyadenylylation of the somatolactin gene in flounder pituitary. By primer extension analysis, the 5′ end of somatolactin mRNA was extended at least 14 bases upstream from the 5′ end of cfSL3 (Fig. 1 *Lower*) (unpublished data). Somatolactin mRNAs were specifically expressed in pituitary but not in the stomach, liver, heart, intestine, spleen, or brain (Fig. 4*B*).

Genomic DNA prepared from a female flounder liver was analyzed by Southern hybridization in order to study somatolactin gene organization. Flounder somatolactin cDNA probes detected a 10.5-kb *Eco*RI fragment and four *Pst* I fragments generated from flounder DNA (Fig. 4*C*). The intensity of the 10.5-kb band indicated the somatolactin gene to be present at about one copy per haploid flounder genome. One *Eco*RI and three *Pst* I fragments prepared from chum salmon DNA were capable of hybridizing with flounder somatolactin cDNA probes, but the hybridization bands were less intense than those obtained for flounder DNA digests.

DISCUSSION

Recently, a 26-kDa glycoprotein was isolated from Atlantic cod pituitaries and shown to be produced by the PIPAS cells and to be structurally related to GH and PRL (M.R.-W., T.N., B. I. Baker, and H.K., unpublished data). By using antibody developed against the 26-kDa protein, a 28-kDa protein was identified in extracts of flounder pituitaries. Immunostaining of flounder pituitary sections with antiserum to the cod 26-kDa protein resulted in specific reaction with PIPAS cells, thus indicating the 28-kDa protein to derive from these cells.

We report here the cDNA structure of the 28-kDa protein, which was tentatively designated somatolactin in consideration of its resemblance to GH and PRL. cDNA clones were isolated from a flounder cDNA library at a frequency of ≈1%, indicating that somatolactin is expressed to a significant degree in the pituitary. Similar cloning frequency was noted for cDNA clones homologous to flounder somatolactin cDNA from chum salmon and Atlantic cod (unpublished data).

Teleost PIPAS cells are considered to secrete a hormone involved in diverse functions such as adaptation to background (2), calcium (4), acid/base conditions (3), and low osmolarity (5). Lack of consensus with regard to function may be a reflection in part of species differences. Though the function of somatolactin is not known, there is a high possibility that it is secreted as a hormone, since its cDNA structure shows the presence of a signal peptide characteristic of secretory proteins; it is produced by the pituitary, an endocrine organ; and it is a member of the GH/PRL family. Its structural similarity to GH and PRL suggests that ion regulation may be the function of somatolactin. This speculation is in agreement with certain histological studies on PIPAS cells (3–5).

Preliminary data provide evidence that somatolactin may be present in all vertebrate classes. Sequences hybridizable with the flounder somatolactin cDNA probes I and II are present in the genomes of the bullfrog, mouse, rat, and human (unpublished data). The presence of a somatolactin-like protein in mammalian pituitaries is also suggested by the staining of tilapia PIPAS cells with antibody raised against a human PRL preparation (23), and proteins related to GH and PRL have been found in rat pituitary extracts (24). All pituitary hormones present in mammals have been identified in lower vertebrates and thus, a somatolactin-like protein should be present in all vertebrate classes.

We are grateful to M. Kawakami for his encouragement and to A. Ito and M. Yamamoto for conducting the computer analyses. This work was supported by a grant-in-aid from the Fisheries Agency of Japan to M.O. and H.K. and from the Ministry of Education, Culture, and Science of Japan (0143021) to H.K.

1. Baker, B. I. (1979) in *Hormone and Evolution*, ed. Barrington, E. J. W. (Academic, New York), Vol. 2, pp. 643–722.
2. Ball, J. N. & Batten, T. F. C. (1981) *Gen. Comp. Endocrinol.* **44**, 233–248.
3. Wendelaar Bonga, S. E., van der Meij, J. C. A. & Flik, G. (1986) *Cell Tissue Res.* **243**, 609–617.
4. Olivereau, M., Olivereau, J.-M. & Aimar, C. (1981) *Cell Tissue Res.* **214**, 23–31.
5. Olivereau, M., Aimar, C. & Olivereau, J.-M. (1980) *Cell Tissue Res.* **212**, 29–38.
6. van Eys, G. J. J. M., Lowik, C. W. G. M. & Wendelaar Bonga, S. E. (1983) *Gen. Comp. Endocrinol.* **49**, 277–285.

7. Nicoll, C. S., Mayer, G. L. & Russell, S. M. (1986) *Endocr. Rev.* **7**, 169–203.
8. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
9. Ono, M., Wada, C., Oikawa, I., Kawazoe, I. & Kawauchi, H. (1988) *Gene* **71**, 433–438.
10. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
11. Feinberg, A. P. & Vogelstein, B. (1983) *Anal. Biochem.* **132**, 6–13.
12. Kanehisa, M. (1982) *Nucleic Acids Res.* **10**, 183–196.
13. Sokal, R. R. & Sneath, P. H. A. (1963) *Principles of Numerical Taxonomy* (Freeman, San Francisco).
14. Lathe, R. (1985) *J. Mol. Biol.* **183**, 1–12.
15. von Heijne, G. (1983) *Eur. J. Biochem.* **133**, 17–21.
16. Watahiki, M., Yamamoto, M., Yamakawa, M., Tanaka, M. & Nakashima, K. (1989) *J. Biol. Chem.* **264**, 312–316.
17. Sekine, S., Mizukami, T., Nishi, T., Kuwana, Y., Saito, A., Sato, M., Itoh, S. & Kawauchi, H. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 4306–4310.
18. Yasuda, A., Itoh, H. & Kawauchi, H. (1986) *Arch. Biochem. Biophys.* **244**, 528–541.
19. Yasuda, A., Miyazima, K., Kawauchi, H., Peter, R. E., Lin, H.-R., Yamaguchi, K. & Sano, H. (1987) *Gen. Comp. Endocrinol.* **66**, 280–290.
20. Seeburg, P. H. (1982) *DNA* **1**, 239–249.
21. Duckworth, M. L., Kirk, K. L. & Friesen, H. G. (1986) *J. Biol. Chem.* **261**, 10871–10878.
22. Linzer, D. I. H. & Nathans, D. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 4255–4259.
23. Rawdon, B. B. (1979) *Gen. Comp. Endocrinol.* **37**, 374–382.
24. Sinha, Y. N., Jacobsen, B. P. & Lewis, U. J. (1989) *Biochem. Biophys. Res. Commun.* **163**, 386–393.