J. of Aquaculture
Vol. 16(3) : 165-170, 2003

한국양식학회지
Journal of Aquaculture
©Korean Aquaculture Society

Molecular Cloning of Red Seabream, *Pagrus major* Somatolactin cDNA and Its Expression in *Escherichia coli*

Helani Munasinghe, Soon-Mi Koh and Jehee Lee*
Department of Marine Biotechnology, Cheju National University, Jeju 690-756, Korea

Isolation, cloning and sequencing of red seabream (*Pagrus major*) somatolactin (rsbSL) cDNA from pituitary gland revealed an open reading frame of 693 bp coding for a pre-growth hormone of 231 amino acids with a 22 amino acid putative signal peptide. Deduced amino acid sequence showed that there was one possible N-glycosylation site at Asn^{145} and seven Cys residues (Cys^{29}, Cys^{39}, Cys^{66}, Cys^{89}, Cys^{205}, Cys^{222}, Cys^{230}). Except Cys^{66}, others may be involved in disulfide bond formation. The rsbSL presented a 93% amino acid sequence identity with the SL of gilthead seabream (*Sparus aurata*) and contained the conserved hormone domain region. Expression of rsbSL in *E. coli* (BL21) cells and gel analysis revealed a higher molecular weight for rsbSL than expected theoretically, implying posttranslational modifications.

Keywords: Cloning somatolactin cDNA, *Pagrus major*, Expression in *E. coli*, Red seabream

Introduction

Somatolactin (SL), a new member of somatotropin hormone family produced in the pars intermedia of the pituitary gland, has been initially isolated from Atlantic cod *(Gadus morhua)* pituitaries as a 26 kDa glycoprotein and later the corresponding proteins have been identified in several teleost species, but not in other classes of vertebrates (Calduch-Giner et al., 1998). Since its discovery in the early 1990s, its biochemical and molecular features have become increasingly clear, while its physiological significance is poorly understood.

Activated SL cells were seen in spawning salmon (Olivereau and Rand-Weaver, 1994) and rainbow trout transferred from calcium-rich water to low calcium water (Kakizawa et al., 1993). An increase in plasma SL level was observed during sexual maturation in coho salmon (Rand-Weaver and Swanson, 1993). Besides, SL was reported to stimulate production of 11-ketotestosterone and testosterone by testicular fragments, and oestradiol by ovarian follicles in a dose-dependent manner in coho salmon (Planas et al., 1992). It was observed that a cobalt variant of rainbow trout, without SL-producing cells, accumulated abdominal fat (Kaneko et al., 1993). An involvement of SL in energy mobilization was also suggested from the chum salmon during spawning migration (Kakizawa et al., 1995). Plasma SL levels also increased during acute stress in rainbow trout (Rand-Weaver et al., 1993) and fasting in gilthead seabream (Company et al., 2001); however, Kakizawa et al. (1996) demonstrated that it was due to blood acidosis rather than stress. These findings suggested SL to be a multifunctional hormone involved in hypercalcaemic regulation, reproduction, stress response, fat mobilization, acid base balance (Kakizawa et al., 1996) and background adaptation (Zhu and Thomas, 1995). Since the other proposed biological events linked to SL more or less affect blood acid-base status in fish, the involvement in acid-base regulation seemed to be the most probable function of SL. It should be noted, however, that most physiological studies on SL were made on salmonids, whose SL is not glycosylated. The physiological function of glycosylated SL in non-salmonids might be different from that of nonglycosylated SL in salmonids. Therefore, the characterization of non-salmonid SL may be crucial in identification of the function of SL in different physiological conditions as well as in recognition of the effect of glycosylation on its activity.

In the present study, the cDNA sequence of SL of a Perciformes fish, red seabream, was cloned, sequenced and expressed in *E. coli*, and was compared with other known SLs or related hormone sequences to identify conserved residues that may help in elucidating its structure-function relationship.

*Corresponding author: jehee@cheju.ac.kr

Materials and Methods

Isolation of mRNA and synthesis of first strand cDNA

Pituitary gland was taken from a red seabream, frozen immediately in liquid nitrogen and stored at -70°C until used. mRNA was extracted from 50 ng of pituitary gland using mRNA Isolation Kit (Invitrogen, USA) following manufacturer's instructions. First-strand cDNA was synthesized using a Double Strand cDNA Synthesis Kit (Promega, USA).

Cloning of the 3' end of red seabream somatolactin cDNA by Rapid Amplification of cDNA Ends (RACE)

A primer SL-F (5'-CTCAACAAGACY(T/C)AAGTGGGT-3') corresponding to a highly conserved region in known teleost SLs, identified by Clustal W 1.71 multiple sequence alignment program, was used with oligo $(dT)_{18}$ primer to amplify the 3' end of rsbSL cDNA. The reaction mixture containing 1×*Pfu* buffer, 0.2 mM each dNTP, 100 ng first strand cDNA, 250 ng of each primer and 2.5 units of *Pfu* DNA polymerase (Stratagene, USA) was processed 30 cycles of amplification with denaturation at 94°C for 30 sec, primer annealing at 53°C for 30 sec and primer extension at 72°C for 1 min with an initial denaturation at 95°C for 1 min and final extension at 72°C for 10 min.

Cloning of the 5' end of red seabream somatolactin cDNA by RACE

A poly(G) tail was added to the 3' end of the first strand cDNA synthesized using the terminal deoxynucleotidyl transferase (Amersham Pharmacia Biotech, UK). This was used as the template to clone the 5' end of the rsbSL cDNA using the same conditions used to amplify the 3' end except primer extension, which was at 72°C for 30 sec with the primers SL-R (5'-ACCCACTTR(G/A) GTCTTGTTGAG-3') and oligo $(dC)_{18}$.

DNA sequencing

PCR products were electrophoresed on 1% agarose gel and purified from agarose using a QIAEX II Gel Extraction Kit (Qiagen, Germany). These were ligated to the pBluescript II SK (-) vector at the *Hin*cII site and transformed into *E. coli* DH 5α competent cells. Plasmid DNA was purified by AccuPrep™ Plasmid Extraction Kit (Bioneer, Korea) and DNA sequencing was executed by ABI 377 DNA sequencer with Dye Terminator Cycle Sequencing Kit.

Expression of red seabream somatolactin cDNA in *Escherichia coli*

rsbSL coding region was amplified from first-strand cDNA by PCR using forward primer PMSL-N (5'-GAGACATATG-CACATGATGAGAGCCATAAAG-3') containing *Nde*1 site and reverse primer PMSL-C (5'-GAGAGGATCCTTA TGCGCAGCTGTATATGTCATTTTG-3') containing *Bam*H1 site. The PCR product digested with the respective enzymes was cloned into pET11a vector (Novagen, USA). This was transformed into BL21 (DE3) cells and recombinant clones were isolated using ampicillin. A recombinant clone was inoculated into 5 ml LB broth containing ampicillin and grew at 37°C overnight. A new LB broth was inoculated with 10% of this overnight grown culture and incubated at 37°C. When cells reached an OD_{600} of 0.5, culture was induced with 1.0 mM IPTG (isopropyl-D-thiogalactopyranoside) and allowed to grow for 3 hrs at 30°C. The cells were harvested and suspended in a buffer containing 50 mM Tris-HCl (pH 8.0), 1 mM EDTA. The cells were lysed with 2x Laemmli sample buffer. The lysate was run on a 12% SDS polyacrylamide gel and stained with Coomassie Brilliant Blue.

Analysis of nucleotide and amino acid sequences

The N-terminal signal peptide of pre-SL was predicted through SignalP World Wide Web server (http://www.cbs.dtu.dk/services/SignalP/), which incorporates a prediction of cleavage sites and a signal peptide/non-signal peptide prediction based on a combination of several artificial neural networks. Prediction of structure was made by Entrez's 3D viewer, which detects related structures by sequence similarity analysis and allows the visualization of sequence-structure alignments. rsbSL sequence was compared with other known SL sequences using CLUSTAL W Multiple Sequence Alignment Program (version 1.8, 1999) and all the other analysis were made using BLAST program (Basic Local Alignment Search Tool), which is a set of programs developed by the National Center for Biotechnology Information (NCBI) to perform similarity searches on all available sequence data.

Results

Isolation, cloning and sequencing of the rsbSL cDNA revealed a coding region of 693 nucleotides, with a 26 bp 5' untranslated region and an 879 bp 3' untranslated region, coding for a 231 amino acid preprotein with a 22 amino acid putative signal peptide region, predicted by the SignalP V2.0 World Wide Web Server (Nielsen et al., 1997a,b) (Fig. 1).

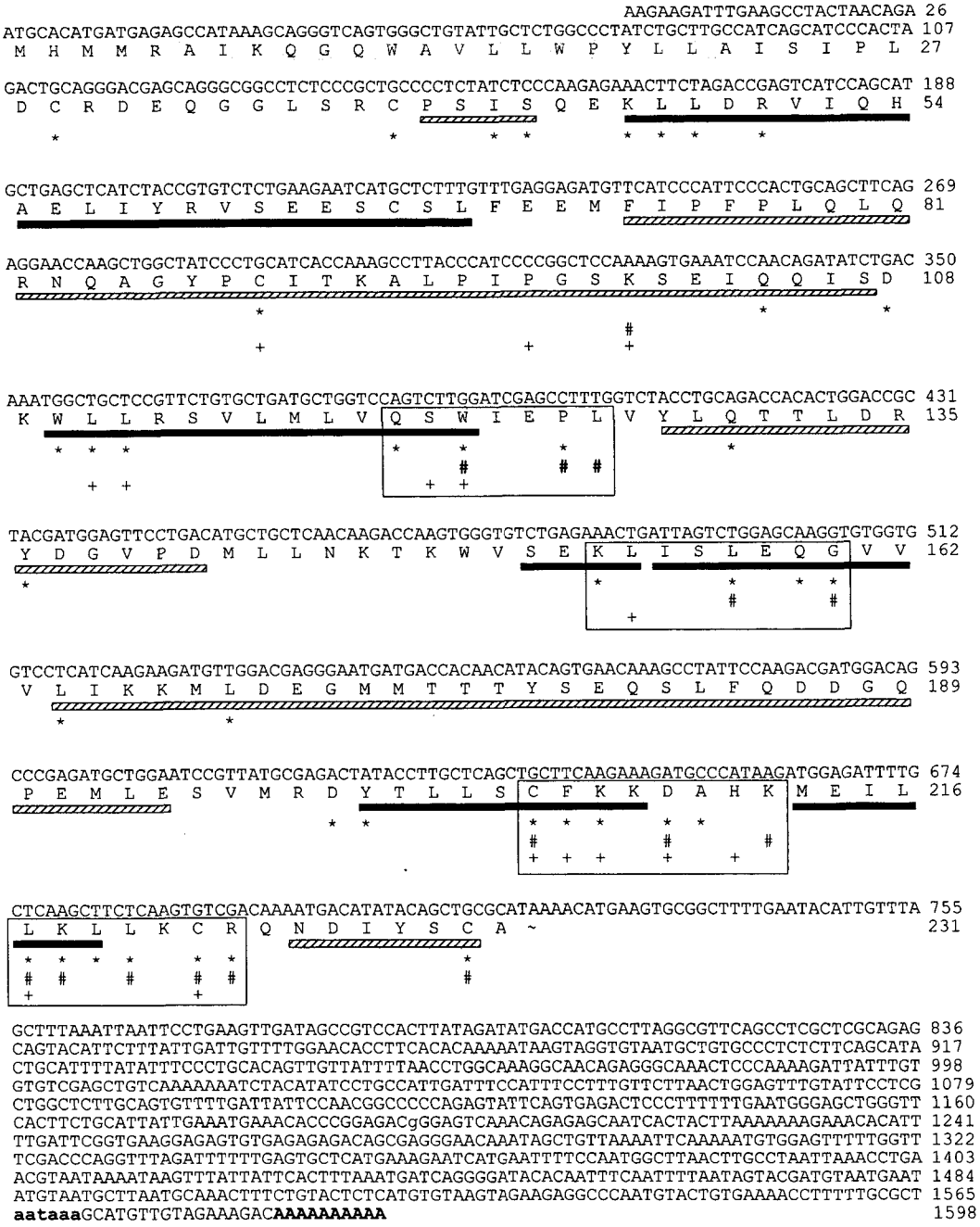

Fig. 1. Nucleotide sequence and the putative amino acid sequence of red seabream somatolactin cDNA. The poly(A) tail is highlighted and the polyadenylation signal is indicated in bold simple case letters. The putative signal peptide sequence is highlighted in ash colour and the evolutionary conserved four domains found in all SLs, prolactins and growth hormones analyzed are boxed. Regions aligned to helixes and extended strands in both human growth hormone and prolactin are indicated by black filled and striped boxes, respectively. Amino acid residues conserved in all 18 known SLs, all 45 known PRLs and 21 growth hormones, selected from 14 different phylogenetic levels ranging from mammals to jawless fish, are marked with asterisks (*), # sign and plus sign (+), respectively.

The prehormone was expected to be 26726.3 Da with a pI of 5.50, containing 14.7% leucine as the major constituent (Expert Protein Analysis System, proteomics server from the Swiss Institute of Bioinformatics).

Comparison of rsbSL with 18 known SL sequences revealed 39 amino acid residues conserved in all SLs but none was in the signal peptide region. These included 6 cysteines (Cys^{29}, Cys^{39}, Cys^{89}, Cys^{205}, Cys^{222}, Cys^{230}) out of the 7 cysteines found in rsbSL (Fig. 2) and 2 (Cys^{205}, Cys^{222}) of these were conserved in all the SLs, growth hormones (GH) and prolactins (PRL) analyzed.

The rsbSL was expressed in *E. coli* (BL21) cells, and anal-

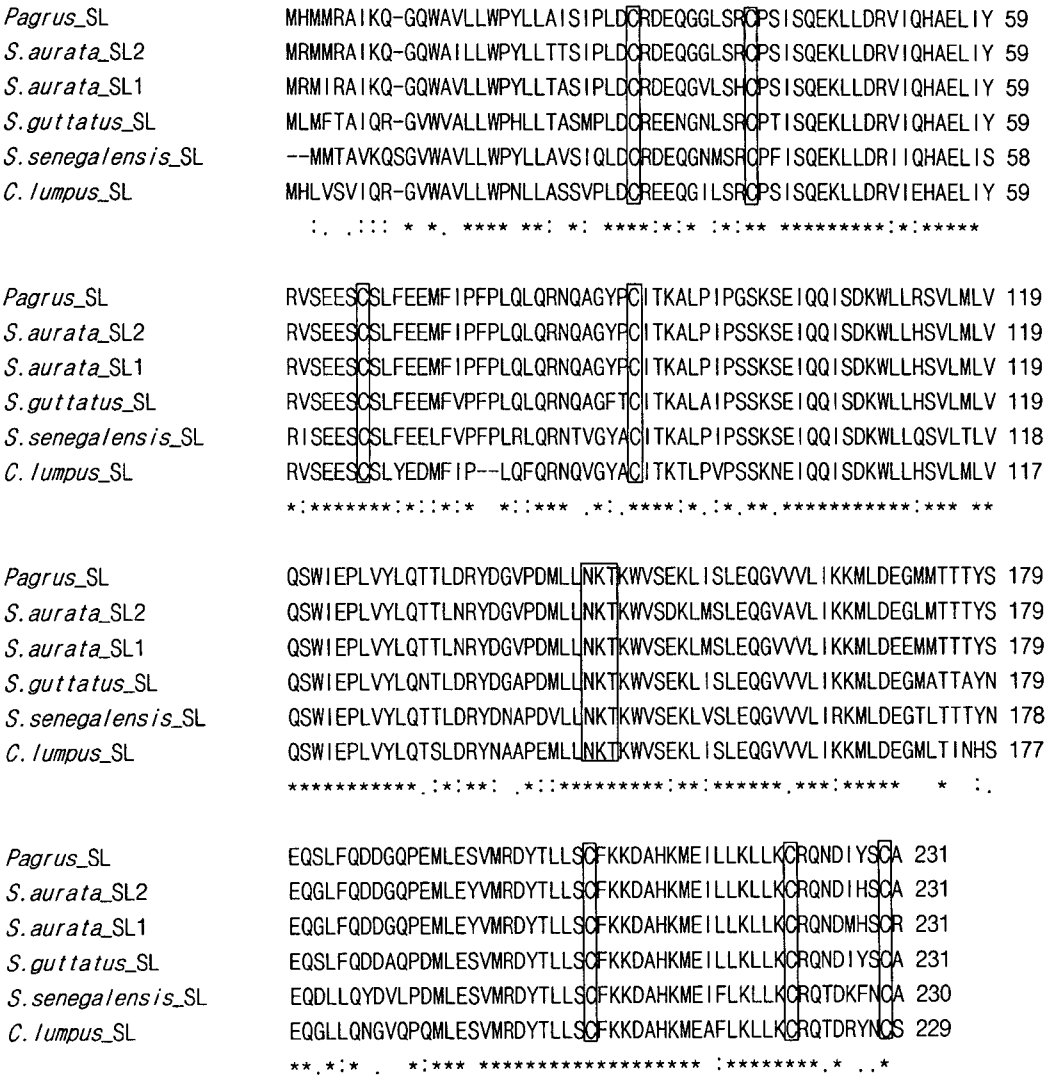

Fig. 2. Multiple sequence alignments of some known somatolactin amino acid sequences. Red seabream SL had score % of 93 with both SL 1 (P54863) and 2 (P79894) of gilthead seabream (*Sparus aurata*). It also showed 87% score with rabbitfish (*Siganus guttatus*, Q9PWG4) and 80% with both Senegalese sole (*Solea senegalensis*, JC2582) and lumpfish (*Cyclopterus lumpus*, P45640) SLs. The cysteine residues are boxed and the putative N-glycosylation site is highlighted. Identical residues are shown in asterisks. Conserved substitutions depending on functionality are indicated with colon and the semi-conserved residues with a dot.

ysis on a gel resulted in a distinct band with a high intensity in IPTG induced samples compared to the samples without IPTG (Fig. 3). The value obtained for preprotein from this was higher than the theoretical one and could be due to post-translational modifications, probably glycosylation at Asn^{145}. The high level expression of rsbSL in *E. coli* suggests the possibility of using this system as a source of rsbSL for future work.

Discussion

Human growth hormone (GH) structure determined by x-ray crystallography has manifested the presence of 2 disulfide bonds between the sites aligned to Cys^{89}-Cys^{205} and Cys^{222}-Cys^{229} of rsbSL sequence in the sequence-structure alignments obtained for rsbSL amino acid sequence and human growth hormone structure by Entrez's 3D viewer, Cn3D. This indicates that there may be a disulfide bond between Cys^{89} and Cys^{205} in rsbSL. Disulfide bond between the sites aligned to Cys^{222} and Cys^{229} produces a short loop on the surface of the human GH molecule. Even though the site corresponding to Cys^{229} of rsbSL in human GH is replaced by a serine in rsbSL, there is a cysteine at position 230 in rsbSL. This can compensate for the loss of Cys^{229} and make a disulfide bond with Cys^{222} producing a larger loop than the one in human GH. Rand-Weaver et al. (1991) have demonstrated the presence of 8 cysteine residues with 3 disulfide bonds in Atlantic

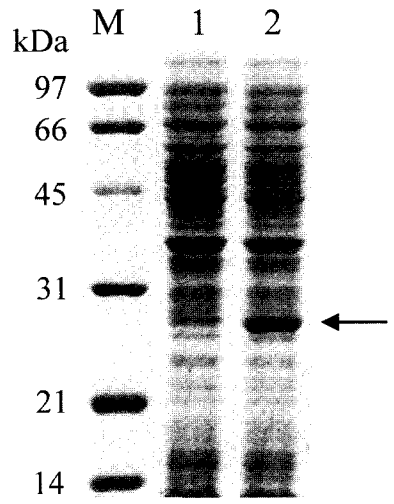

Fig. 3. Analysis of total proteins expressed in *E. coli* (BL21) cells containing red seabream somatolactin coding region in a 12% denaturing polyacrylamide gel. Lane M - SDS-PAGE low range molecular weight standards (BioRad, USA) in kDa, lane 2 and 1 - *E. coli* (BL21) cells containing rsbSL coding region with and without induction. The arrow indicates the position of the band corresponding to SL.

cod. The multiple sequence alignment of SL sequences of the 2 species has revealed the presence of 2 of these disulfide bonds at positions aligned to Cys^{89}-Cys^{205} and Cys^{222}-Cys^{230} supporting the above prediction. The other disulfide bond found in Atlantic cod align to Cys^{29} and Cys^{39} in rsbSL, a region not found in human GH. But there is a disulfide bond in human PRL between positions aligned to Cys^{32} and Cys^{39} making a short loop. The cysteine at position 29 might compensate for the loss of cysteine at position 32 and make the third disulfide bond with Cys^{39} in rsbSL. The other two cysteines found in Atlantic cod aligns to Cys^{66} and Ser^{204} in rsbSL and are not involved in disulfide bond formation.

The alignment of cDNA sequences of all 18 known SLs by CLUSTAL W (1.8) multiple sequence alignment program revealed high conservation at nucleotide level (25%) compared to amino acid level (17%) as calculated by the percentage of number of residues conserved in all SL coding sequences analyzed. Since the number of residues in the nucleotide sequence is 3 times higher than the amino acid sequence it codes, the nucleotide sequence can hold 3 times more mutations than that of the amino acid sequence, while maintaining the same level of conservation. Since the mutations in the nucleotide sequence have a higher chance of being in codons coding for amino acids at different positions than on the same amino acid, this may lead to a lower percent conservation at amino acid level. The chance of a mutation in the amino acid sequence to be deleterious is higher than a mutation in the nucleotide sequence. Since most of the mutations have adverse effects rather than being useful, the selective pressures will act more on amino acid level than on nucleotide level to conserve sequences. The above observation may be a balance between these two opposing forces. The absence of conserved sequences at the untranslated regions suggest that natural selection acts with different strengths on untranslated regions, coding regions and amino acid level.

From the sequence-structure alignments obtained for rsbSL amino acid sequence with human GH and human prolactin (PRL) structures by Entrez's 3D viewer, Cn3D, we have identified regions where both hormones have similar secondary structures. We expect rsbSL also to have similar secondary structures at the same positions. In human PRL, all of these helices are buried inside the molecule and probably involved in structure formation than having any functional role unless they are exposed due to a structural change. But in GH, the residues in helices are more exposed and the first two helices are on the surface of the molecule. In both, the extended strands are generally on the surface. Therefore, the extended strands in SL are also expected to be on the surface and may play a functional role and the helices may be mainly involved in structure stabilization.

Multiple alignments of cDNA sequences of 18 SLs, 21 growth hormones and 45 prolactins showed the presence of 4 domains highly conserved in all 3 hormones. They were mainly found inside of human GH and PRL. They may be the remnants of the ancestral gene(s), from which all the three hormones were derived and could have formed the foundation for the development of somatotropin hormone family. SL is a new member identified in this family and its production in substantial quantity is essential for further analysis.

Acknowledgement

Soon-Mi Koh was supported by a research assistant program from the Brain Korea 21 Project.

References

Calduch-Giner, J. A., C. Pendon, M. M. Valdivia and J. Perez-Sanchez, 1998. Recombinant somatolactin as a stable and bioactive protein in a cell culture bioassay: development and validation of a sensitive and reproducible radioimmunoassay. J. Endocrinol., **156**: 441–447.

Company, R., A. Astola, C. Pendon, M. M. Valdivia and J. Perez-Sanchez, 2001. Somatotropic regulation of fish growth and

adiposity: growth hormone (GH) and somatolactin (SL) relationship. Comp. Biochem. Physiol. C, **130**: 435–445.

Kakizawa, S., T. Kaneko, S. Hasegawa and T. Hirano, 1993. Activation of somatolactin cells in the pituitary of the rainbow trout *Oncorhynchus mykiss* by low environmental calcium. Gen. Comp. Endocrinol., **91**: 298–306.

Kakizawa, S., T. Kaneko, S. Hasegawa and T. Hirano, 1995. Effects of feeding, fasting, background adaptation, acute stress and exhaustive exercise on the plasma somatolactin concentrations in rainbow trout. Gen. Comp. Endocrinol., **98**: 137–146.

Kakizawa, S., T. Kaneko and T. Hirano, 1996. Elevation of plasma somatolactin concentrations during acidosis in rainbow trout. J. Exp. Biol., **199**: 1043–1051.

Kaneko, T., S. Kakizawa and T. Yada, 1993. Pituitary of "cobalt" variant of the rainbow trout separated from the hypothalamus lacks most pars intermedial and neurohypophysial tissue. Gen. Comp. Endocrinol., **92**: 31–40.

Nielsen, H., J. Engelbrecht, S. Brunak and G. von Heijne, 1997a. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Engineering, **10**: 1–6.

Nielsen, H., J. Engelbrecht, S. Brunak and G. von Heijne, 1997b. A neural network method for identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Int. J. Neural Sys., **8**: 581–599.

Olivereau, M. and M. Rand-Weaver, 1994. Immunocytochemical study of the somatolactin cells in the pituitary of Pacific salmon, *Oncorhynchus nerka* and *O. keta* at some stages of the reproductive cycle. Gen. Comp. Endocrinol., **93**: 28–35.

Planas, J. V., P. Swanson, M. Rand-Weaver and W. W. Dickhoff, 1992. Somatolactin stimulates *in vitro* gonadal steroidogenesis in coho salmon, *Oncorhynchus kisutch*. Gen. Comp. Endocrinol., **87**: 1–5.

Rand-Weaver, M., T. Noso, K. Muramoto and H. Kawauchi, 1991. Isolation and characterization of somatolactin, a new protein related to growth hormone and prolactin from Atlantic cod (*Gadus morhua*) pituitary glands. Biochemistry, **30**: 1509–1515.

Rand-Weaver, M. and P. Swanson, 1993. Plasma somatolactin levels in coho salmon (*Oncorhynchus kisutch*) during smoltification and sexual maturation. Fish. Physiol. Biochem., **11**: 175–182.

Rand-Weaver, M., T. G. Pottinger and J. P. Sumpter, 1993. Plasma somatolactin concentrations in salmonid fish are elevated by stress. J. Endocrinol., **138**: 509–515.

Zhu, Y. and P. Thomas, 1995. Red drum somatolactin: Development of a homologous radioimmunoassay and plasma levels after exposure to stressors or various backgrounds. Gen. Comp. Endocrinol., **99**: 275–288.

Manuscript Received: June 23, 2003
Revision Accepted: July 14, 2003
Responsible Editorial Member: Jong-Man Yoon