Proc. Natl. Acad. Sci. USA
Vol. 82, pp. 4306–4310, July 1985
Applied Biology

Cloning and expression of cDNA for salmon growth hormone in *Escherichia coli*

(nucleotide/amino acid sequences/bacterial expression/growth-promoting activity)

SUSUMU SEKINE*, TAMIO MIZUKAMI*, TATSUNARI NISHI*, YOSHIHISA KUWANA*, AKIKO SAITO*, MORIYUKI SATO*, SEIGA ITOH*†, AND HIROSHI KAWAUCHI‡

*Tokyo Research Laboratories, Kyowa Hakko Kogyo Co., Ltd. 3-6-6, Asahimachi, Machidashi, Tokyo 194, Japan; ‡School of Fisheries Sciences, Kitasato University, Sanriku, Iwate 022-01, Japan

Communicated by John E. Halver, March 5, 1985

ABSTRACT **cDNA clones encoding chum salmon (*Oncorhynchus keta*) growth hormone (sGH) have been isolated from a cDNA library prepared from chum salmon pituitary gland poly(A)$^+$ RNA. Synthetic oligodeoxynucleotide mixtures based on amino acid residues 23–28 of sGH were used as hybridization probes to select recombinant plasmids carrying the sGH coding sequence. The complete nucleotide sequence of sGH cDNA has been determined. The cDNA sequence codes for a polypeptide of 210 amino acids, including a putative signal sequence of 22 amino acids. The 5′ and 3′ untranslated regions of the message were 64 and 426 bases long, respectively. Mature sGH was efficiently expressed in *Escherichia coli* carrying a plasmid in which the sGH cDNA was under control of the *E. coli trp* promoter; sGH comprised about 15% of the total cellular protein in such bacteria. The partially purified sGH from *E. coli* stimulated the growth of rainbow trout and the activity was indistinguishable from that of natural sGH.**

Human growth hormone now can be produced by genetically engineered organisms and can be used as a therapeutic agent. Salmon growth hormone (sGH) can be synthesized by use of similar techniques, and the massive supply of sGH may be extremely important to fish culture. Growth hormone (GH), together with prolactin and chorionic somatomammotropin (placental lactogen), forms a set of proteins that are structurally related and have partially overlapping biological activities (1). Primary structure analysis of the peptides and of the genes suggests that these hormone genes evolved from a common ancestral origin (1–5). Therefore, these genes provide an excellent model system for studying structure–function relationships, evolution, and regulation of expression. GH genes have been isolated from several mammalian species and characterized in detail (3, 6–8). To obtain information about the evolution and the mechanisms of organization of this set of genes, it is essential to compare the structures of these hormone genes isolated from many organisms at various evolutionary stages. No information, however, has been available about lower vertebrates such as fish.

Recently, growth hormones from tilapia (9), sturgeon (10), carp (11), and chum salmon (unpublished result) and prolactin from chum salmon (12) were purified. In the case of chum salmon growth hormone, a partial amino acid sequence has been determined (unpublished data). Based on this partial amino acid sequence, oligodeoxynucleotides complementary to the putative sGH mRNA were synthesized and used as probes to identify cDNA clones corresponding to the sGH mRNA in a salmon pituitary gland cDNA library. Here we describe the isolation and characterization of a cDNA encoding sGH. Expression of the cDNA in *Escherichia coli* gave rise to a polypeptide having growth-promoting activity equivalent to that of authentic sGH prepared from salmon pituitary gland.

```
                 1   23  24  25  26  27  28
amino acid     NH2-Ile---Lys Met Phe Asn Asp Phe---

Probe A             5'-AAA ATG TTT AAT GAT TT-3'
                         G       C   C   C


                    167 168 169 170 171   188
amino acid          ---His Lys Val Glu Thr---Leu-COOH

Probe B             5'-CAT AAA GTT GAA AC-3'
                         C   G   C   G
                                 A
                                 G
```

FIG. 1. Partial amino acid sequence of sGH near the NH_2 terminus (residues 23–28) and near the COOH terminus (residues 167–171) and synthetic oligodeoxynucleotide probes corresponding to those amino acid sequences.

MATERIALS AND METHODS

Preparation of a cDNA Library. Poly(A)$^+$ RNA from chum salmon pituitaries was prepared by the guanidinium/cesium chloride method (13), followed by chromatography on oligo(dT)-cellulose. Using this poly(A)$^+$ RNA as template, we prepared a cDNA library with the plasmid vectors developed by Okayama and Berg (14) and *E. coli* K-12 C600 SF8 as a host strain.

Screening of the cDNA Library. Two oligodeoxynucleotide mixtures (Fig. 1) were synthesized by the solid-phase phosphotriester method (15). These probes were complementary to the NH_2-terminal region (probe A) and COOH-terminal region (probe B) of the putative sGH mRNA sequence predicted from the partial amino acid sequence of sGH.

First, probe A was used to detect colonies containing sGH cDNA by the method of Wallace *et al.* (16). Hybridizing colonies were grown in liquid culture, and plasmid DNAs were extracted and digested with restriction enzymes, followed by Southern blot hybridization (17) with probe A and probe B for confirmation and further characterization.

DNA Sequence Analysis. The sequence of the cloned sGH cDNA was determined by the method of Maxam and Gilbert (18) and by the dideoxynucleotide chain-termination method (19) after subcloning fragments in the M13 phage vectors (20, 21).

Expression of sGH in *E. coli*. *E. coli* K-12 W3110 (*strA*) containing psGHIB-2, a plasmid for expression for sGH (see

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.

Abbreviations: GH, growth hormone; sGH, salmon GH; bp, base pair(s).
†To whom reprint requests should be addressed.

Results), was cultured in M9/Casamino acid medium at 30°C for 16–18 hr. Cells were collected, dissolved in Laemmli sample buffer (22) directly, and analyzed by $NaDodSO_4$/PAGE (22) with Coomassie brilliant blue staining.

Assay for Biological Activity of sGH. A crude sample of sGH produced in *E. coli* was assayed for its effect on the growth rate of rainbow trout less than 1 year old [9–9.5 cm fork length (front tip of fish to end of backbone), 10–12 g body weight]. The fish were kept in running water at 8–10°C and maintained on a 12-hr light/12-hr dark photoperiod before and during the experiment. All fish were individually marked. The hormone was administered by intraperitoneal injection once a week. Dosage of the hormone was 1 μg/100 μl of 0.9% NaCl per fish. Control fish received 100 μl of 0.9% NaCl. Fish were fed to satiation with pellets (Masu No. 4P; Nihonhaigo Shiryo) twice daily, and weight and fork length of fish were measured at each time of injection.

RESULTS

Isolation of cDNA for sGH Gene. Polyadenylylated RNA was prepared from chum salmon pituitaries and used for cDNA synthesis according to the method of Okayama and Berg (14). The resulting cDNA library was screened by colony hybridization with ^{32}P-labeled probe A (Fig. 1) complementary to the NH_2-terminal region of the putative sGH mRNA sequence predicted from the partial amino acid sequence of the protein. Twenty-eight clones were obtained from 4800 transformants, and their plasmids were cleaved with several restriction endonucleases and analyzed by Southern blot hybridization (17) with probe A and probe B (Fig. 1); 8 clones with cDNA inserts of 1.0–1.3 kilobase pairs (kbp) were found to hybridize with both probes. From restriction endonuclease analysis, these clones were divided into two groups: recombinant plasmids psGH-1, -3, -6, -9, -10, and -17 contained *Bgl* II, *Pvu* II (two sites), *Sal* I, *Kpn* I, *Stu* I, *Sac* I, and *Hin*dIII sites (Fig. 2A), whereas psGH-8 and psGH-14 carried *Bgl* II, *Pvu* II (one site), and *Sac* I sites but no *Kpn* I, *Sal* I, *Stu* I, or *Hin*dIII sites (data not shown). As probes A and B hybridized to these two types of clones, it is likely that these two types of cDNA code for polypeptides that are homologous, at least around the portion corresponding to the probes. In addition, restriction endonuclease cleavage sites for *Pvu* II (5′-proximal site in the psGH-1 type), *Bgl* II, and *Sac* I were located at similar positions, respectively, in both psGH-1- and psGH-14-type plasmids, suggesting that these two types of cDNA resulted from different mRNA species that were transcribed from independent genes that differ from each other by some small deletions or substitutions.

Nucleotide and Derived Amino Acid Sequence of sGH cDNA. Because 6 clones of the psGH-1 type and only 2 clones of the

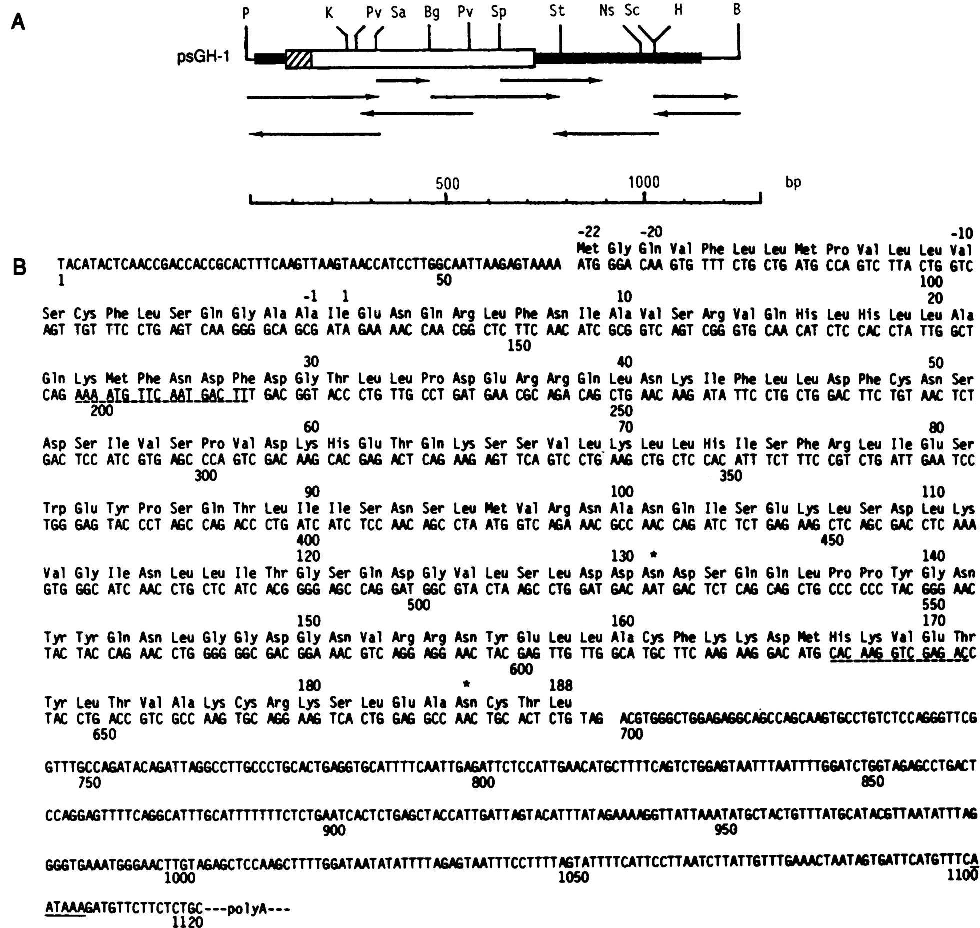

FIG. 2. (*A*) Restriction map of the cDNA insert of psGH-1. From left to right, the diagram shows a schematic representation of the G·C tail (thin line), the 5′ untranslated region (thick line), the regions coding for the signal peptide (hatched box) and mature sGH (open box), the 3′ untranslated region (thick line), and the A·T tail (thin line). The sequencing strategy is indicated under the map; arrows indicate direction and extent of sequencing. Restriction enzyme cleavage sites are indicated as follows: B, *Bam*HI; Bg, *Bgl* II; H, *Hin*dIII; K, *Kpn* I; Ns, *Nsi* I; P, *Pst* I; Pv, *Pvu* II; Sa, *Sal* I; Sc, *Sac* I; Sp, *Sph* I; St, *Stu* I. (*B*) The complete nucleotide sequence and deduced amino acid sequence of psGH-1. Pre-sGH contains the signal peptide (amino acids −22 to −1) and the mature protein (amino acids 1–188). Potential N-glycosylation sites (Asn-Xaa-Ser and Asn-Xaa-Thr) are marked by asterisks, and the polyadenylylation signal AATAAA (nucleotides 1100–1105) is underlined. The portions corresponding to the probes are indicated by dashed lines (nucleotides 197–213 for probe A and 629–642 for probe B).

psGH-14 type were obtained, it is likely that the mRNA complementary to the psGH-1 type may predominate among the mRNAs that contain the sequences homologous to probes A and B. Accordingly, we determined the complete nucleotide sequence of the cDNA insert from psGH-1 by the dideoxynucleotide chain-termination method (19–21) and the chemical-cleavage method (18), according to the strategy illustrated in Fig. 2*A*.

The sequence of the psGH-1 insert (Fig. 2*B*) contains a single large open reading frame of 630 nucleotides. The amino acid sequence predicted by this region is given above the nucleotide sequence. The amino acid sequences (residues 1–40 and 167–188) are identical with the partial amino acid sequence for sGH (unpublished results). The first ATG, which usually serves as the initiation codon in eukaryotes, is found at nucleotides 65–67 from the 5′ end. This ATG is followed by 210 codons before the termination triplet TAG (positions 695–697). The 3′ untranslated region of 426 nucleotides contains the hexanucleotide AATAAA (positions 1100–1105) which precedes the polyadenylylation site in many eukaryotic mRNAs (23).

A primary structure of the sGH polypeptide consisting of 188 amino acids could be deduced (Fig. 2*B*), and the molecular weight was calculated to be 21,556. The 22 NH_2-terminal amino acids encoded by the cDNA insert of psGH-1 may constitute a signal peptide that is cleaved off in the secretion process.

Comparison of amino acid sequence homologies between sGH and rat GH (3) or human GH (6) is shown in Fig. 3; gaps were inserted to maximize homologies. In the coding regions of sGH and rat GH, there is 39% and 53% homology for amino acid and nucleotide sequence, respectively. sGH has 35% amino acid and 50% nucleotide homology with human GH. There are two potential N-glycosylation sites (Asn-Xaa-Ser and Asn-Xaa-Thr) in the predicted amino acid sequence of sGH, at Asn-131 and Asn-185.

Codon usage in sGH mRNA is shown in Table 1. The codon usage is rather nonrandom, and there is some preference for G or C over A or U at the third position of codons (75%).

Table 1. Codon usage in sGH mRNA

UUU/Phe	2	UCU/Ser	4	UAU/Tyr	0	UGU/Cys	2
UUC/Phe	7	UCC/Ser	3	UAC/Tyr	6	UGC/Cys	3
UUA/Leu	1	UCA/Ser	2	UAA/Stop	0	UGA/Stop	0
UUG/Leu	4	UCG/Ser	0	UAG/Stop	1	UGG/Trp	1
CUU/Leu	0	CCU/Pro	2	CAU/His	1	CGU/Arg	1
CUC/Leu	6	CCC/Pro	2	CAC/His	4	CGC/Arg	1
CUA/Leu	2	CCA/Pro	2	CAA/Gln	4	CGA/Arg	0
CUG/Leu	20	CCG/Pro	0	CAG/Gln	9	CGG/Arg	2
AUU/Ile	2	ACU/Thr	2	AAU/Asn	2	AGU/Ser	4
AUC/Ile	7	ACC/Thr	4	AAC/Asn	13	AGC/Ser	6
AUA/Ile	2	ACA/Thr	0	AAA/Lys	2	AGA/Arg	2
AUG/Met	5	ACG/Thr	1	AAG/Lys	10	AGG/Arg	3
GUU/Val	0	GCU/Ala	1	GAU/Asp	3	GGU/Gly	1
GUC/Val	9	GCC/Ala	3	GAC/Asp	10	GGC/Gly	3
GUA/Val	1	GCA/Ala	2	GAA/Glu	3	GGA/Gly	2
GUG/Val	4	GCG/Ala	2	GAG/Glu	6	GGG/Gly	4

Numbers indicate the frequency with which the codons are used in the coding region of sGH mRNA.

Expression of sGH in *E. coli*. A partial amino acid sequence for sGH purified from salmon pituitary glands has been determined (unpublished data), but the primary structure of sGH was not completely elucidated. Therefore, to prove that the cDNA sequence of psGH-1 actually encodes sGH, this cloned gene was expressed in *E. coli* and the biological activity of the product was measured.

The *E. coli* system was selected for expression of the cDNA, because it has been used to produce human GH efficiently (24). The strategy for construction of a plasmid for sGH expression, illustrated in Fig. 4, is analogous to that we have used for production of human interferon γ. This construction involves a synthetic DNA linker, corresponding to the NH_2-terminal portion of sGH and shown in Fig. 4. This approach allows the direct expression of the mature hormone

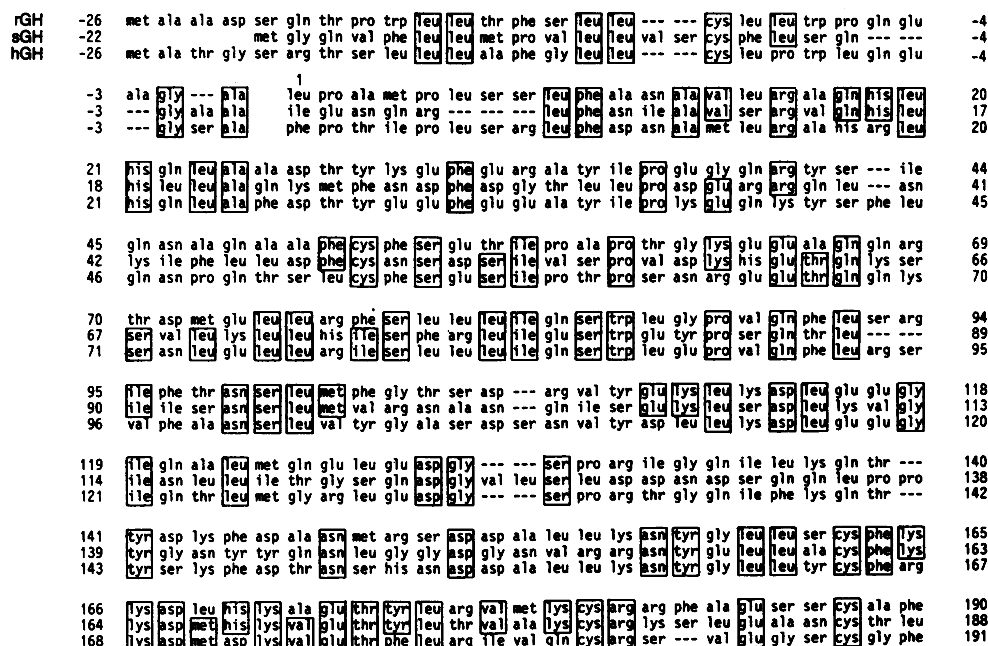

FIG. 3. Comparison of amino acid sequences of preprocessed sGH, rat GH (rGH), and human GH (hGH). The amino acid sequences were aligned by introducing gaps to maximize homology. Amino acid residues identical to pre-sGH are boxed. The NH_2-terminal amino acids of mature hormones (Ile for sGH, Leu for rGH, Phe for hGH) are taken as position 1; the amino acids in signal peptides are given negative numbers.

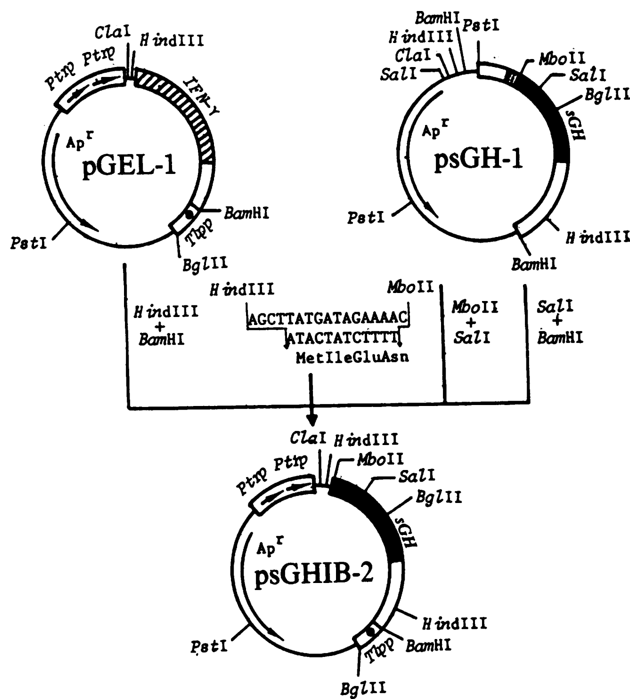

FIG. 4. Schematic representation of construction of an expression plasmid for sGH. Plasmid pGEL-1 consists of tandemly linked *trp* promoter, human interferon γ (IFN-γ) cDNA, *lpp* terminator, and a segment from pBR322 containing an origin of replication and ampicillin resistance (Ap^r) gene. psGH-1 is a donor plasmid of the sGH gene which was inserted in the cloning vector of Okayama and Berg (14). A large *Hin*dIII–*Bam*HI fragment of pGEL-1 was isolated, which contains segments required for a vector function and efficient gene expression. A 163-bp *Mbo* II–*Sal* I fragment of psGH-1, which encodes a small NH_2-terminal portion of sGH, and an ≈1-kbp *Sal* I–*Bam*HI fragment, which carries the remaining portion of sGH cDNA, were prepared independently. A DNA linker was synthesized for combining the *trp* promoter and sGH gene segment and for introducing an additional ATG codon just before the triplet (for Ile) encoding the amino-terminal residue of mature-form sGH. These three fragments and the synthetic linker DNA were ligated to construct a plasmid for expression of sGH, designated psGHIB-2. The nucleotide sequence between the Shine–Dalgarno sequence (AAGG) and the ATG codon in plasmid psGHIB-2 is 14 bp long (AAGGGTATCGATAAGCTTATG). *Ptrp*, tryptophan operon promoter (the direction of transcription is indicated by an arrow); *Tlpp*, lipoprotein terminator.

by introducing an ATG initiation codon in front of the first amino acid codon of the mature sGH. Plasmid pGEL-1 contains the *trp* promoter, human interferon γ gene, and *lpp* terminator in adjacent positions to maximize the expression of interferon γ (details will be described elsewhere). Actually, *E. coli* cells harboring pGEL-1 synthesize a large amount of interferon γ, >20% of total cellular proteins. Two fragments required for expression of the mature sGH were prepared independently; a 163-bp *Mbo* II–*Sal* I fragment of psGH-1 that encodes the NH_2-terminal region of sGH and an ≈1-kbp *Sal* I–*Bam*HI fragment that carries the remaining portion of sGH cDNA. These two fragments and synthetic DNA linker were assembled and inserted between the *Hin*dIII and *Bam*HI sites of pGEL-1. The resulting plasmid, psGHIB-2, contains the mature sGH coding sequence and 3′ noncoding region just downstream of the tandem *trp* promoter and *lpp* transcriptional terminator downstream of the sGH cDNA. psGHIB-2 was introduced into *E. coli* strain W3110 (*strA*), and the resulting transformants were cultured and then analyzed by $NaDodSO_4$/PAGE as described in *Materials and Methods*. A distinct band corresponding to a protein of about 25 kDa was visible which was absent from extracts of bacteria containing a vector plasmid only. Since this expression system should constitutively synthesize sGH because of the titration of repressor molecules by the multiple copies of the *trp* operator, a procedure for induction of the promoter was omitted. Although the amount of sGH synthesized in *E. coli* carrying psGHIB-2 was estimated to comprise about 15% of total cellular protein, we expect the yield can be increased by modification of the culture conditions or the plasmid. In addition, inclusion bodies were observed in the *E. coli* cells containing psGHIB-2, and the particles were easily recovered and purified from sonicated cells by centrifugation. The particles were analyzed by $NaDodSO_4$/PAGE, which revealed a major band at about 25 kDa, which is identical in size to the protein band observed in whole cell extracts of psGHIB-2-harboring bacteria. As sGH prepared from salmon pituitary glands was only slightly soluble in water, sGH synthesized in bacterial cells was also thought to aggregate to form particles. The particles collected by centrifugation were solubilized in 8 M urea solution and renatured by dilution and dialysis as described by Marston *et al.* (25), with a slight modification. The partially purified sGH sample was used in a growth-promoting experiment described in the legend to Fig. 5 and in *Materials and Methods*. The results clearly indicate that sGH synthesized in *E. coli* is equipotent to the natural sGH in promoting increases in weight and length of rainbow trout.

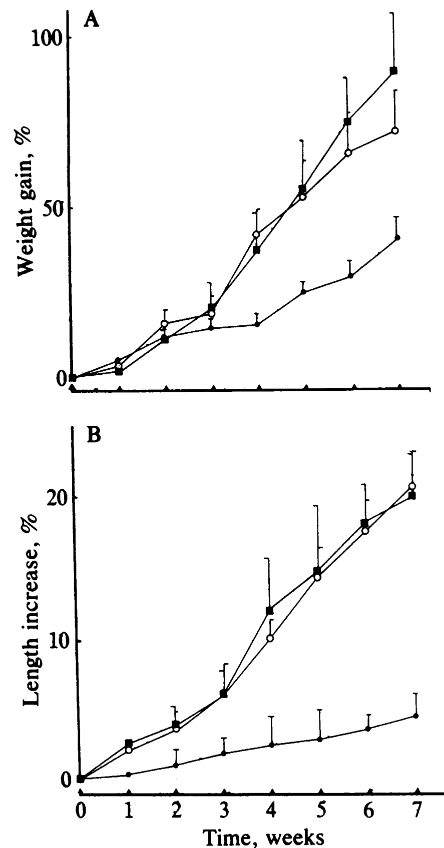

FIG. 5. sGH produced in *E. coli* (■, n = 8 fish per group) and natural sGH (○, n = 10) both stimulate growth in weight (*A*) and length (*B*) of rainbow trout less than 1 year old. ●, Control (saline-injected) fish (n = 10). Vertical lines represent +SEM.

DISCUSSION

In this paper, we have described the cloning and expression of sGH in *E. coli*. The sGH protein sequence deduced from the nucleotide sequence of a cloned sGH cDNA is in good agreement with the partial amino acid sequence for sGH (unpublished results).

The nucleotide sequence predicts the complete amino acid and signal peptide sequence of sGH, which had not been determined directly. Comparison of the DNA and deduced amino acid sequence for sGH with mammalian GHs permitted us to make some speculations about evolutionary changes between fish and mammals. The putative signal peptide of sGH is hydrophobic like other signal peptides and is 22 amino acids long, which is shorter than that of human or rat GH (26 amino acids) or bovine GH (27 amino acids). The deduced amino acid sequence of the signal peptide shows low overall homology to the signal peptide sequence of human (35%) or rat GH (42%), but it is noteworthy that leucines at sGH positions −11, −12, −16, and −17, cysteine at −8, and glycine at −3 are located in the homologous positions in all three GHs.

The deduced sequence of mature sGH shows 39% and 35% homology with rat and human GH, respectively. Although the extent of homology is rather low, some unique structural features are found among these GHs. First, there are four cysteine residues at nearly identical positions in all three GHs. Cysteine residues can form disulfide bonds, which affect conformation of protein molecules. These four cysteine residues of sGH should form disulfide bonds in a fashion similar to rat GH or human GH. Since cysteine residues are located at nearly identical positions in bovine, porcine, equine, and ovine GHs (26), it is likely that these residues play an essential role in maintaining the biologically active form of GH. Second, a highly conserved region exists near the COOH terminus (Asn-155 to Arg-179), whose homology to rat GH is 76% (19 of 25 residues) and to human GH, 64% (16 of 25 residues). Although conserved amino acid sequences are also observed in other regions, the degree of homology is relatively low.

This highly conserved region near the COOH terminus, containing three cysteine residues, may be important for biological activity, but the active fragment required for the GH activity remains to be identified. A structural feature characteristic of sGH is the presence of two potential N-glycosylation sites (Asn-Xaa-Ser or Asn-Xaa-Thr). No N-glycosylation site has been found in the mammalian growth hormones, and sGH isolated from salmon pituitary glands was also considered to be nonglycosylated (unpublished observation). As the presence of these sequences is a necessary but not sufficent condition for glycosylation, it will be interesting to see whether sGH can be glycosylated in cells of other eukaryotes, such as yeasts or animals.

sGH produced in *E. coli* containing psGHIB-2 was estimated to be about 15% of the total cellular protein, and the protein formed inclusion bodies in cells. There are some cases of soluble proteins that are expressed in *E. coli* in an insoluble form (27). It is a great advantage to obtain proteins produced in *E. coli* as insoluble granules because it is possible to sediment them with a single centrifugation step after disruption of cells, and the contaminating impurities such as cell membrane proteins can be eliminated by washing the pellet with mild detergents. After denaturation and renaturation of the granule protein, soluble sGH was obtained which showed a significant biological activity on the growth of rainbow trout. The effect of the recombinant sGH on the growth rate of the fish was indistinguishable from that of natural sGH.

As the cloning and expression of sGH would facilitate the production of larger quantities of this growth hormone, it should be possible to test whether sGH may have a commercial value in fish culture. Moreover, by *in vitro* mutagenesis, it should be possible to dissect the active site(s) of the molecule, because the growth-promoting activity of sGH on fish can be examined more easily than those of other mammalian growth hormones.

We are grateful to Drs. Takeo Suzuki, Akira Furuya, and Tetsuo Oka for encouragement and support throughout this work. We also thank Miss Takako Amoh and Miss Masako Odaka for excellent technical assistance and Dr. Fusao Tomita for critical reading of the manuscript.

1. Niall, H. D., Hogan, M. L., Tregar, G. W., Segre, G. V., Hwang, P. & Friesen, H. (1973) *Recent Prog. Horm. Res.* **29**, 386–416.
2. Niall, H. D., Hogan, M. L., Sayer, R., Rosenblum, I. Y. & Greenwood, F. C. (1971) *Proc. Natl. Acad. Sci. USA* **68**, 866–869.
3. Seeburg, P. H., Shine, J., Martial, J. A., Baxter, J. D. & Goodman, H. M. (1977) *Nature (London)* **270**, 486–494.
4. Shine, J., Seeburg, P. H., Martial, J. A., Baxter, J. D. & Goodman, H. M. (1977) *Nature (London)* **270**, 494–499.
5. Cooke, N. E., Coit, D., Weiner, R. I., Baxter, J. D. & Martial, J. A. (1980) *J. Biol. Chem.* **255**, 6502–6510.
6. Martial, J. A., Hallewell, R. A., Baxter, J. D. & Goodman, H. M. (1979) *Science* **205**, 602–607.
7. Miller, W. L., Martial, J. A. & Baxter, J. D. (1980) *J. Biol. Chem.* **255**, 7521–7524.
8. Seeburg, P. H., Sias, S., Adelman, J., Boer, H. A., Hayflick, J., Thurani, P., Goeddel, D. V. & Heyneker, H. L. (1983) *DNA* **2**, 37–45.
9. Farmer, S. W., Papkoff, H., Hayashida, T., Bewley, T. A., Bern, H. A. & Li, C. H. (1976) *Gen. Comp. Endocrinol.* **30**, 91–100.
10. Farmer, S. W., Hayashida, T., Papkoff, H. & Polenov, A. L. (1981) *Endocrinology* **108**, 377–381.
11. Cook, A. F., Wilson, S. W. & Peter, R. E. (1983) *Gen. Comp. Endocrinol.* **50**, 335–347.
12. Kawauchi, H., Abe, K., Takahashi, A., Hirano, T., Hasegawa, S., Naito, N. & Nakai, Y. (1983) *Gen. Comp. Endocrinol.* **49**, 446–458.
13. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), p. 196.
14. Okayama, H. & Berg, P. (1983) *Mol. Cell. Biol.* **3**, 280–289.
15. Ito, M., Ike, Y., Ikura, S. & Itakura, K. (1982) *Nucleic Acids Res.* **10**, 1755–1769.
16. Wallace, R. B., Johnson, M. J., Hirose, T., Miyake, T., Kawashima, E. H. & Itakura, K. (1981) *Nucleic Acids Res.* **9**, 879–894.
17. Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503–517.
18. Maxam, A. M. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 560–564.
19. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
20. Messing, J. & Vieira, J. (1982) *Gene* **19**, 269–276.
21. Messing, J., Crea, R. & Seeburg, P. H. (1981) *Nucleic Acids Res.* **9**, 309–321.
22. Laemmli, U. K. (1970) *Nature (London)* **227**, 680–685.
23. Proudfoot, N. J. & Brownlee, G. G. (1976) *Nature (London)* **263**, 211–214.
24. Goeddel, D. V., Heyneker, H. L., Hozumi, T., Arentzen, R., Itakura, K., Yansura, D. G., Ross, M. J., Miozzari, G., Crea, R. & Seeburg, P. H. (1979) *Nature (London)* **281**, 544–548.
25. Marston, F. A. O., Lowe, P. A., Doel, M. T., Schoemaker, J. M., White, S. & Angal, S. (1984) *Bio/Technology* **2**, 800–804.
26. Dayhoff, M. O. (1976) *Atlas of Protein Sequence and Structure* (National Biochemical Research Foundation, Washington, DC), Vol. 5, Suppl. 2, pp. 120–121.
27. Harris, T. J. R. (1983) in *Genetic Engineering*, ed. Williamson, R. (Academic, London), Vol. 4, pp. 127–183.