

## Sequence Analysis of the Protein Structure Homology Modeling of Growth Hormone Gene from *Salmo trutta caspius*

<sup>1</sup>Abolhasan Rezaei and <sup>2</sup>Sheyda Akhshabi

<sup>1</sup>Department of Genetics-School of Basic Science, Islamic Azad University  
Tonekabon Branch, Iran

<sup>2</sup>Young Researchers Club, Tonekabon Branch, Islamic Azad University, Tonekabon, Iran

**Abstract:** In view of the growth hormone protein investigated and characterized from *Salmo trutta caspius*. Growth hormone gene in the *Salmo trutta caspius* have six exons in the full length that is translated into a Molecular Weight (kDa): ssDNA: 64.98 and dsDNA: 129.6. There are also 210 amino acid residue. The assembled full length of DNA contains open reading frame of growth hormone gene that contains 15 sequences in the full length. The average GC content is 47% and AT content is 53%. This protein multiple alignment has shown that this peptide is 100% identical to the corresponding homologous protein in the growth hormone protein which including *Salmo salar* (Accession number: AAA49558.1) and *Rainbow trout (Salmo trutta)* (Accession number: AAA49555.1") sequences. The sequence of protein had deposited in Gene Bank, Accession number: AEK70940. Also we were analyzed second and third structure between sequences reported in Gene Bank Network system. The results are shown, there are homology between second structure in three sequences including: *Salmo trutta caspius*, *Salmo salar* and *Rainbow trout*. Regarding third structure, *Salmo trutta caspius* and *Salmo salar* are same type, but *Rainbow trout* has different homology with *Salmo trutta caspius* and *Salmo salar*. However, the sequences were observed three parallel  $\alpha$  helix and in second structure there were almost same percent  $\beta$  sheet.

**Key words:** Growth hormone protein, *Salmo trutta caspius*

### INTRODUCTION

The salmonids comprise eleven genera that include salmon, trout, charr, freshwater whitefishes, ciscos and graylings (Nelson, 2006). Many salmonid species are of considerable economic, social and environmental importance (Gross and Nilsson, 1999). Salmonids contribute to local and global economies through aquaculture, wild stock fisheries and recreational sport fisheries. There are so much gene regarding to increase production and accelerates growth rate in the fisheries. Growth hormone gene is unique gene in the fisheries. This gene as a potential growth and metabolism function. Growth hormone controls of the postnatal somatic growth (Reiter and Rosenfeld, 1998). And reproduction (Van der Kraak *et al.*, 1990; Le Gac *et al.*, 1993). In vertebrates, these functions of GH are mediated by insulin like growth factors IGF 1 and 2, (Mathews *et al.*, 1986; Sara, 1991) and growth related protein kinases (Ralph *et al.*, 1990). The growth hormone gene in salmonids sequenced and reported in Gene Bank Network system. That including fifteen sequences from full length of GH gene. However

almost all the described GH sequences are for fishes of Europe and western countries. The GH genes in Salmons have two types, type 1 and 2. There are several studies regard GH genes, have demonstrated that GH genes are highly variable in Salmonid fishes (Forbes *et al.*, 1994). Studied genetic variation within GH1 and GH2 in *Brown trout*, *Salmo salar* and other Salmonid fishes and found length polymorphism in amplification products of both genes (Gross and Nilsson, 1995). The results also are shown; there are high homology between sequences of GH genes in the Salmons. The Iranian salmon namely *Salmo trutta caspius*. Constitute the most important cultivable freshwater fishes of Iran. In this study DNA of GH have been sequenced that including 2048 bp. in the length and including 210 amino acids residues. The site of encoding amino acids is from 28 to 208. The aligned with other sequences reported in Gene Bank, results are shown the site of encoding amino acid started from 28 to 208. Moreover an attempt has also been made for the first time to describe the second and 3-D structure of GH gene in the *Salmo trutta caspius*.

**MATERIALS AND METHODS**

**Samples:** The samples of *Salmo trutta caspius* including Bloods and Muscles has taken from three old age females. The Salmons originated from Rivers of Tonekabon-Iran.

**DNA preparation:** Genomic DNA was isolated from blood samples following the phenol-chloroform extraction method described by Sambrook *et al.* (1989). DNA was dissolved in TE buffer and was kept in a water bath at 60°C for 2 h to dissolve pellet properly in buffer. The quality of DNA was checked through spectrophotometry. DNA samples with O.D. ratio between 1.7 and 1.9 were considered as good and used for further study. The samples beyond this range were re-extracted by the phenol-chloroform extraction method. DNA quality was also checked by running the sample in 1.5% agarose gel electrophoresis. The DNA samples devoid of smear were used for further study.

**DNA amplification:** A 2048 bp full length of the growth hormone gene spanning over six exon and five was amplified. Forward primer: (5'ACATACTCAA CCGACC ACCGCACTTTCAAG-3') and Reverse primer: (5' TCAGTGCAGGGCAAGGCCTAATCTG 3') primer sequences. PCR was carried out in a final volume of 50 µL reaction mixture containing 80-100 ng DNA, 5 µL, 10X ChromTaq Assay buffer, 2 µL dNTP mix (2.5 mM each), 2.0 µL Forward primer (100 ng/µL), 2 µL Reverse primer (100 ng/µL), 2 µL. dNTP mix (2.5 mM each). 0.5

µL. ChromTaq enzyme (3 U/µL). 37.5 µL. water. PCR-reactions were carried out in a thermocycler (PTC-200, MJ Research, USA) in two stages. In the first stage, initial denaturation at 94°C for 5 min, denaturation at 94°C for 30 sec, annealing at 55°C for 30 sec, extension at 72°C for 1 min were carried out for 35 cycles.

**Sequencing of DNA and protein GH:** The PCR products were concentrated to 50 ng/µL by pooling several tubes to precipitate by the isopropanol procedure. In order to obtain clean fragment for sequencing, the PCR products were separated by electrophoresis in a TAE agarose gel containing ethidium bromide using standard protocols. The desired PCR product band was excised using a clean, sterile razor blade or scalpel (band was visualized in a medium or long wavelength (e.g., ≥300 nm) UV light, and excised quickly to minimize exposure of the DNA to UV light). The minimum agarose slice was transferred to a 1.5 mL microcentrifuge or screw cap tube and then purified by using commercially available gel extraction kits (Qiagen). Quantification was done by loading one µL of eluted sample in 1.5% agarose gel and comparing with standard molecular marker (Phi X 174 DNA ladder or 500 bp DNA ladder). Only samples with good concentration (>50 ng/µL) were selected and subjected to sequencing.

**The analysis of sequence:** Sequence data were analyzed mostly by DNAMAN software and also conducted using the BLAST Network system with other sequences of GH reported in Gene Bank.

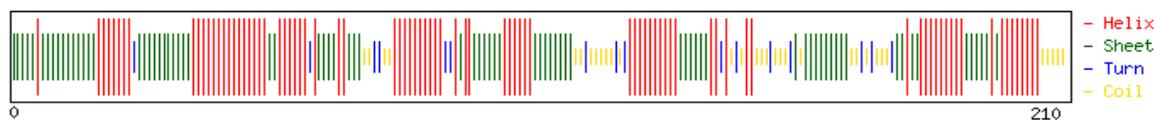
S.t.caspius	MGQVFLLM	PVLLVSCF	LSQG	20
S.salar	MGQVFLLM	PVLLVSCF	LSQG	20
R.trutta	MGQVFLLM	PVLLVSCF	LSQG	20
Consensus	mgqvfllm	pvllvscf	lsqg	
S.t.caspius	AA	ENQRLFNI	AVNRVQHLH	40
S.salar	AA	ENQRLFNI	AVNRVQHLH	40
R.trutta	AA	ENQRLFNI	AVNRVQHLH	40
Consensus	aa	enqrlfni	avrvqhlh	
S.t.caspius	LAQKMFNDF	GTLL	DERR	60
S.salar	LAQKMFNDF	GTLL	DERR	60
R.trutta	LAQKMFNDF	GTLL	DERR	60
Consensus	l	aqkmfndf	gtll derr	
S.t.caspius	QLNKIFLLDF	CNSDSIVSP		80
S.salar	QLNKIFLLDF	CNSDSIVSP		80
R.trutta	QLNKIFLLDF	CNSDSIVSP		80
Consensus	qlnkiflldf	cnsdsivsp		
S.t.caspius	DK	ETQKSSVLKLLHISFRL		100
S.salar	DK	ETQKSSVLKLLHISFRL		100
R.trutta	DK	ETQKSSVLKLLHISFRL		100
Consensus	dk	etqkssv	klhhisfrl	

MGQVFLLM PV LLVSCF LSQG AAIENQRLFNI IAVSRVQHLH LLAQKMFNDF DGTLLPDERR QLNKIFLLDF  
 CNSDSIVSPV DKHETQKSSV LKLLHISFRL IESWEYPSQT LIISNSLMVR NANQISEKLS DLKVGINLLI  
 TGSQDGVLSL DDNDSQQLPP YGNYYQNLGG DGNVRRNYEL LACFKKDMHK VETYLTVAKC RKSLEANCTL

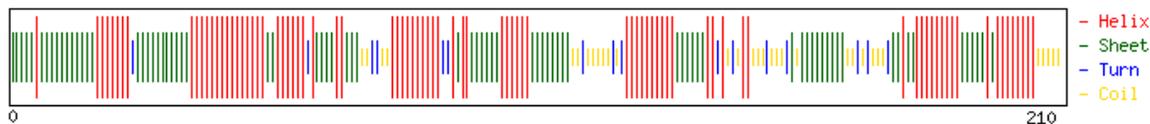
Fig. 1: Target Sequence: Sequence consists of 210 amino acids of GH in *Salmo trutta caspius*, including: five cysteine residue that started from amino acid of 14. The sequence deposited in GeneBank, Accession number: AEK70940.1

S . t . caspius	I E S W E Y P S Q T L L I S N S L M V R	1 2 0
S . salar	I E S W E Y P S Q T L L I S N S L M V R	1 2 0
R . trutta	I E S W E Y P S Q T L L I S N S L M V R	1 2 0
Consensus	i e s w e y p s q t l i s n s l m v r	
S . t . caspius	N S N Q I S E K L S D L K V G I N L L I	1 4 0
S . salar	N S N Q I S E K L S D L K V G I N L L I	1 4 0
R . trutta	N S N Q I S E K L S D L K V G I N L L I	1 4 0
Consensus	n n q i s e k l s d l k v g i n l l i	
S . t . caspius	R G S Q D G V L S L D D N D S Q L P P	1 6 0
S . salar	R G S Q D G V L S L D D N D S Q L P P	1 6 0
R . trutta	T G S Q D G V L S L D D N D S Q L P P	1 6 0
Consensus	g s q d g v l s l d d n d s q l p p	
S . t . caspius	Y G N Y Y Q N L G G D G N F R R N Y E L	1 8 0
S . salar	Y G N Y Y Q N L G G D G N F R R N Y E L	1 8 0
R . trutta	Y G N Y Y Q N L G G D G N F R R N Y E L	1 8 0
Consensus	y g n y y q n l g g d g n f r r n y e l	
S . t . caspius	L A C F K K D M H K V E T Y L T V A K C	2 0 0
S . salar	L A C F K K D M H K V E T Y L T V A K C	2 0 0
R . trutta	L A C F K K D M H K V E T Y L T V A K C	2 0 0
Consensus	l a c f k k d m h k v e t y l t v a k c	

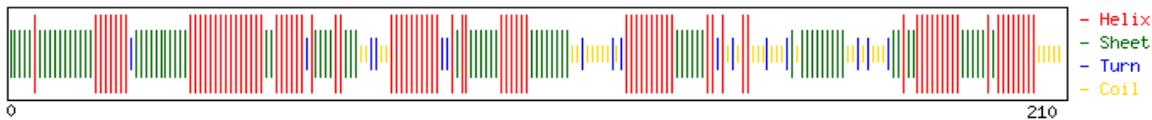
Fig. 2: Comparison of the GH protein amplification products in *Salmo trutta caspius*, *Salmo salar* and *Rainbow trout*. The results are shown; there is high homology between sequences



(A) *Salmo trutta caspius* : Total Residues: H: 146 E: 109 T: 27 Percent: H: 69.5 E: 51.9 T: 12



(B) *Salmo salar* : Total Residues: H: 148 E: 105 T: 29 Percent: H: 70.5 E: 50.0 T: 13.8



(C) *Rainbow trout*: Total Residues: H: 152 E: 110 T: 26 Percent: H: 72.4 E: 52.4 T: 12.4

Fig. 3: Secondary structure prediction of the GH protein of *Salmo trutta caspius*, (A) *Salmo salar*, (B) and *Rainbow trout*, (C) Protein sequence of GH is annotated with secondary structure information,  $\alpha$ -helix in red, (h);  $\beta$  sheet in green, (e); random coil in yellow, (c), Turn in blue (B). The *Salmo trutta caspius*, *Salmo salar* and *Rainbow trout* GH proteins have almost the same structure and the  $\alpha$ -helix to  $\beta$  sheet ratio is almost constant in all these Salmons

Deposited of sequence: The sequence of protein deposited in NCBI Network, Accession number: AEK70940.

## RESULTS

**Nucleotide sequencing:** Polymorphism of the genotype of growth hormone gene was confirmed by nucleotide sequencing. The nucleotide sequence as well as the derived amino acid sequence of growth hormone gene

of *Salmo trutta caspius* (Gene Bank Accession Number AEK70940.1) (Fig. 1). Nucleotide sequence variations and amino acid variations were observed at the growth hormone protein between *Salmo trutta caspius*, *Salmo salar* and *Rainbow trout* reveals 100% nucleotide identity (Fig. 2).

**Second and third structure of the GH protein on *Salmo trutta caspius*:** The third type of structure found in

proteins is called tertiary protein structure. The tertiary structure is the final specific geometric type that a protein assumes. This final type is determined by a variety of bonding interactions between the "side chains" on the amino acids. These bonding interactions may be stronger than the hydrogen bonds between amide groups holding the helical structure. As a result, bonding interactions between "side chains" may cause a number of folds, bends, and loops in the protein chain. Different fragments of the same chain may become bonded together. In this study, there were three structures from second and 3-D structure of GH protein's *Salmo trutta caspius*, *Salmo salar* and *Rainbow trout*. The results observed there were high homology regard second structure between three

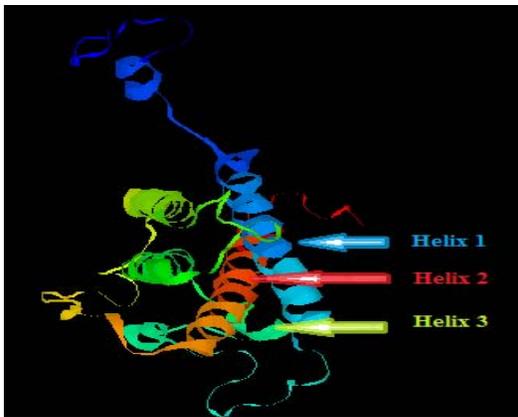


Fig. 4: 3-D structure of the *Salmo trutta caspius* GH. Almost 150, amino acids (69.5%) are involved in  $\alpha$ -helix formation, which run anti-parallel to each other. These helices attribute to the typical 3- $\alpha$ -helix bundle protein conformation, the characteristic 3-D confirmation of growth hormones

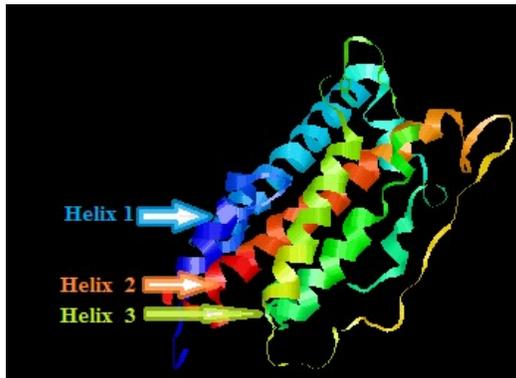


Fig. 5: 3-D structure of the *Salmo salar* GH. Almost 151, amino acids (70.5%) are involved in  $\alpha$ -helix formation, which run anti-parallel to each other. These helices attribute to the typical 3- $\alpha$ -helix bundle protein conformation, the characteristic 3-D confirmation of growth hormones

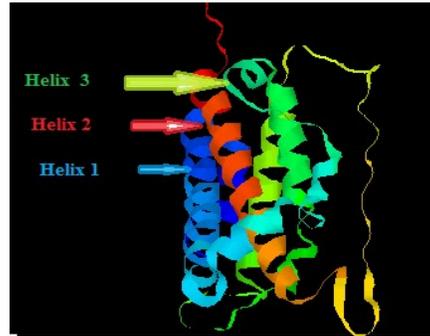


Fig. 6: 3-D structure of the *Rainbow trout* GH. Almost 153, amino acids (72.5%) are involved in  $\alpha$ -helix formation. These helices attribute to the typical 3- $\alpha$ -helix bundle protein conformation, the characteristic 3-D confirmation of growth hormones

sequences, however, the percent of  $\alpha$  Helix (H), 69.5, 70.5 and 72.4,  $\beta$  Sheet (E): 51.9, 50 and 52.4. and, Turn (T): 12.9, 13.8 and 12.4, *Salmo trutta caspius*, *Salmo salar* and *Rainbow trout*, respectively. But regard 3-D structure, there were same type between *Salmo trutta caspius* and *Salmo salar* GH protein, but regard *Rainbow trout* was different types. (Fig. 3, 4, 5 and 6).

## DISCUSSION

**Gene duplication in salmon:** Following the discovery of GH genes in the bonyfishes (Male *et al.*, 1992; Rezaei *et al.*, 2011a, b), the hypothesis of an ancestral genome duplication in bony fishes gained rapid popularity (Postlethwait *et al.*, 1998; Holland, 1999; Meyer and Schartl, 1999). If this "more genes in fish" theory (Wittbrodt *et al.*, 1998) fits with the empirical experience of Zebrafish genetics who often notice two copies of their favorite gene where only one was described in mice, it lacks solid backing. Indeed, GH gene complexes only give information on two loci in vertebrates, and cannot be assumed to represent the whole genome. Moreover, to attribute a duplication to an evolutionary lineage, such as bony fish, a phylogenetic analysis specifying the order of events (speciations and duplications) is necessary. GH genes are very good phylogenetic markers, (Gross and Nilsson, 1995), however the GH gene relevant to paternal traits more than maternal traits. The GH gene sequenced from *Salmo trutta caspius* in this study encodes for the GH protein. It has a start codon from the nucleotide number of in the same translation frame, and the deduced polypeptide has a length of 210 aa. That start codon started from amino acid of 28<sup>th</sup> to 218<sup>th</sup> that included: 28,31,34..35,37,40..41,44,60..61,64..65,67,79..82,85..86,131,134..135,138,182,185..186,189..190,192..193,196..197,200,208. (Fig. 7).

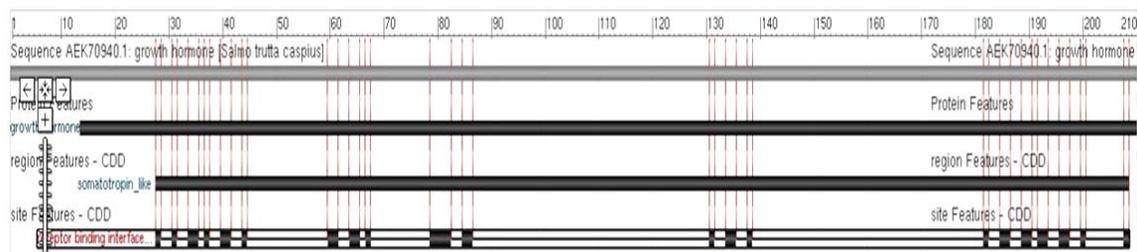


Fig. 7: Structure of graphical growth hormone gene in the *Salmo trutta caspius* (Accession number: JN241634.1). Exons are labeled, Exon1, Exon 2, Exon 3, Exon 4, Exon 5 and Exon 6. Size is indicated one base pairs (bp). Positions and directions of exons are shown by color arrows

Table 1: The percent of amino acid sequence of GH protein in *salmo trutta caspius*

Ala	Cys	Asp	Glu	Phe	Gly	His	Ile	Lys	Leu	MET	Asn	Pro	Gln	Arg	Ser	Thr	Val	Trp	Tyr
8	5	13	9	9	10	5	11	12	33	5	15	6	13	9	19	7	14	1	6

Regarding to *Salmo salar* GH gene also the regions of site had same homology. But about *Rainbow trout* was not same homology; however between the sequences had high homology.

The third structure of protein of GH in *salmo trutta caspius* and other salmons: In this study compared the protein structures of the GH gene from *Salmo trutta caspius* with *Salmo salar* and *Rainbow trout*. 3D-structure of the *Salmo trutta caspius* GH protein is some variation with *Salmo salar* and *Rainbow trout*. In related to, in all the known crystallographic and theoretical GH structures, more than 70% of the residues are involved in the formation  $\alpha$ -helices in anti-parallel twisted helical bundle. Sequence comparison indicates that the  $\alpha$ -helical regions are more conserved than other parts of the molecule (Wallis, 1989), suggesting the possibility of similar conformation in all the other proteins of the GH family. Members of the GH family possess either two (GH) or three (PRL) disulfide bonds, which are highly conserved.

The GH protein of all the vertebrates species contain four Cysteine residues that render the functional 3-D confirmation of the hormone by establishing two intramolecular disulfide (S-S) bonds. As an interesting exception, the GH of all fish species studied to date specially, the family 'Salmonids' including *Salmo trutta caspius*, *Salmo salar* and *Rainbow trout*, contain five Cysteine residues. The recent, (Vestling *et al.*, 1991) reported that in the vertebrates, the first disulfide 15-61-183-190-208 bond is formed between Cysteine 49 and Cysteine 161, and the second between Cysteine 178 and Cysteine 186. In addition to the four residues, which occur at positions similar to those of other vertebrates (Chang *et al.*, 1992) while in the fish, GH protein, first disulfide bonding from amino acid of Cysteine will be started 15<sup>th</sup> to 61, and second disulfide bonding started from 183 to 190 and other Cysteine located in position 208. The results also showed that there are regions of rich

amino acid of Leucine that multiple of 15 amino acid repeats, while Methionine and Cysteine had same number repeat, that regard other vertebrates was different (Table 1). Protein Length = 210. Predicted pI = 6.59.

In comparison with mammalian GH proteins, GH proteins in fish contain a polar amino acid at different position, which is replaced by a nonpolar amino acid. For instance, the GH of Salmon contain 19 Serine and 15 Asparagine at different positions of the length GH gene, but the human GH contains a nonpolar (Leu) residue at this positions. Moreover, with to do sequence of GH amino acid in *Salmo trutta caspius*, homology among the GH proteins of fishes suggests that the GH proteins of other fishes might share a similar 3-D structure. However, the construction of 3-D protein in GH of *Salmo trutta caspius* is similar with *Salmo salar* but regarding *Rainbow trout* 3-D structure is different, however there is homology between together sequences of amino acids. Also, was reported that, there are two glycosylation site in the GH protein which produce different forms of the GH protein. However the GH secreted in the pituitary gland (hGH) of human is not glycosylated but that of the placenta during pregnancy is glycosylated. In fish, glycosylation of the GH protein is reported in salmon (Wagner *et al.*, 1985), but whether there is a corresponding glycosylation in Salmonids specially *Salmo trutta caspius* is not yet known.

#### ACKNOWLEDGMENT

This study had financially supported by the research council of Islamic Azad University Tonekabon branch.

#### REFERENCES

- Chang, Y.S., C.S. Liu, F.L. Huang and T.B. Lo, 1992. The primary structures of growth hormones of three cyprinoids species: Bighead carp, Silver carp and grass carp. Gen. Comp. Endocrinol., 87: 385-393.

- Forbes, S.H., K.L. Knudsen, T.W. North and F.W. Allendorf, 1994. One of two growth hormone genes in coho salmon is sex-linked. Proc. Natl. Acad. Sci. U.S.A., 91: 1628-1631.
- Gross, R. and J. Nilsson, 1995. Application of heteroduplex analysis for detecting variation within the growth hormone 2 gene in (*Salmo trutta* L.) brown trout. Hered., 74: 286-295.
- Gross, R. and J. Nilsson, 1999. Restriction fragment length polymorphism at the growth hormone 1 gene in Atlantic salmon (*Salmo salar* L.) and its association with weight among the offspring of a hatchery stock. Aquaculture, 173: 73-80.
- Holland, P.W., 1999. Gene duplication: Past, present and future. Cell Dev. Biol., 10: 541-547.
- Le Gac, F., O. Blaise, A. Fostier, P.Y. Le Bail, M. Loir, B. Mourot and C. Weil, 1993. Growth Hormone (GH) and reproduction: A review. Fish Phys. Biochem., 11: 219-232.
- Male, R., A.H. Nerland, J.B. Lorens, W. Telle, I. Lossius and G.K. Totland, 1992. The complete nucleotide sequence of Atlantic salmon growth hormone I gene. Biochim. Biophys. Acta., 1130: 345-348.
- Mathews, L., G. Norstedt and R.D. Palmiter, 1986. Regulation of insulin like growth factor I gene expression by growth hormone. Proc. Natl. Acad. Sci. USA., 83: 9343-9347.
- Meyer, A. and M. Schartl, 1999. Gene and genome duplications in vertebrates: The one-to-four (-to-eight in fish) rule and the evolution of novel gene functions. Curr. Opin. Cell. Biol., 1: 699-704.
- Nelson, J.S., 2006. Fishes of the World. 3rd Edn., Wiley and Son Davidson, William, New York.
- Ralph, R.K., S. Darkin-Rattray and P. Schofield, 1990. Growth related protein kinases. Bio. Essays, 12: 121-124.
- Reiter, E.O. and Rosenfeld, 1998. Normal and Abnormal Growth. In: Sakamoto, T., S.D. McCormick and T. Hirano, 1993. (Eds.), Williams Text Book of Endocrinology. 9th Edn., Osmoregulatory Actions of Growth Hormone and its Mode of Action in Salmonids: A Review. Fish. Physiol. Biochem., 11: 155-164.
- Rezaei, A., S.H. Akhshabi and H.R. Jamalzadeh, 2011a. Investigation and Identification *salmo trutta caspius* by the Growth Hormone type 1 (GH1) gene analogue. B. J. Biotech., 3: 76-82.
- Rezaei, A., S.H. Akhshabi and H.R. Jamalzadeh, 2011b. Sequencing and sequence analysis of Growth Hormone type 1 (GH1) gene homologue in the *salmo trutta caspius*. Adv. S. Biol., 3: 111-122.
- Sambrook, J., E.F. Fritsch and T. Maniatis, 1989. Molecular Cloning: A Laboratory Manual. 2nd Edn., Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
- Sara, V.R., 1991. Insulin like growth factor and fetal growth. Encycl. Hum. Biol., 4: 503-511.
- Postlethwait, J.H., Y.L. Yan, M.A. Gates, S. Horne, A. Amores, A. Brownlie, A. Donovan, E.S. Egan, A. Force and Z. Gong, 1998. Vertebrate genome evolution and the zebrafish gene map. Nat. Genet., 18: 345-349.
- Van der Kraak, G., P.M. Rosenblum and R.E. Peter, 1990. Growth hormone dependent potentiation of gonadotropin stimulated steroid production by ovarian follicles of the gold fish. Gen. Comp. Endocrinol., 71: 233-239.
- Vestling, M., C. Murphy, C. Fenselau and T.T. Chen, 1991. Disulfide bonds in native and recombinant fish growth hormone. Mol. Mar. Biol. Biotechnol., 1: 73-77.
- Wallis, M., 1989. Species Specificity and Structure-Function Relationships of Growth Hormone. In: Heap, R.B., C.G. Prosser and G.E. Lamming, (Eds.), Biotechnology in Growth Regulation, Butterworths, London, pp: 3-14.
- Wagner, G.M., R.C. Fargher, J.C. Brown and B.A. McKeown, 1985. Further characterization of growth hormone from the chum salmon (*Oncorhynchus keta*). Gen. Comp. Endocrinol., 1: 27-34.
- Wittbrodt, J., A. Meyer and M. Schartl, 1998. More genes in fish? Bioessays, 20: 511-515.