Title:

Circulating salmon 41-kDa insulin-like growth factor binding protein (IGFBP) is not IGFBP-3 but an IGFBP-2 subtype

Authors:

Munetaka Shimizu[a,b]*, Seira Suzuki[a], Moeri Horikoshi[a], Akihiko Hara[a], and Walton W. Dickhoff[b]

Affiliations

[a]Faculty of Fisheries Sciences, Hokkaido University, 3-1-1 Minato, Hakodate, Hokkaido 041-8611, Japan

[b]Northwest Fisheries Science Center, NOAA Fisheries, 2725 Montlake Blvd. E., Seattle, WA 98112; and School of Aquatic and Fishery Sciences, University of Washington, Seattle, WA 98195, USA

*Corresponding author: Tel&Fax: +81-138-40-8897

E-mail address: mune@fish.hokudai.ac.jp

Abstract

In vertebrates, most circulating insulin-like growth factor (IGF) is bound to multiple forms of IGF-binding proteins (IGFBPs) that differ both structurally and functionally. In mammals, the largest reservoir of IGF in the circulation comes from a large (150kDa) ternary complex comprised of IGF bound to IGFBP-3, which is bound to an acid label subunit (ALS), and this variant of IGFBP is regulated by growth hormone (GH) and feed intake. Although multiple variants of IGFBPs ranging from 20 to 50 kDa have been found in fishes, no ternary complex is present and it has been assumed that the majority of circulating IGF is bound to fish IGFBP-3. Consistent with this assumption is previous work in salmon showing the presence of a 41-kDa IGFBP that is stimulated by GH, decreases with fasting and increases with feeding. However, the hypothesis that the salmon 41-kDa IGFBP is structurally homologous to mammalian IGFBP-3 has not been directly tested. To address this issue, we cloned cDNAs for several Chinook salmon IGFBPs, and found that the cDNA sequence of the 41-kDa IGFBP is most similar to that of mammalian IGFBP-2 and dissimilar to IGFBP-3. We found an additional IGFBP (termed IGFBP-2a) with high homology to mammalian IGFBP-2. These results demonstrate that salmon 41-kDa IGFBP is not IGFBP-3, but a paralog of IGFBP-2 (termed IGFBP-2b). Salmon IGFBP-2s are also unique in terms of having potential N-glycosylation sites and splice variants. Additional research on non-mammalian IGFBPs is needed to fully understand the molecular/functional evolution of the IGFBP family and the significance of the ternary complex in vertebrates.

Keywords

insulin-like growth factor binding protein; salmon; identification; gene duplication; N-glycosylation; splicing variants

Abbreviations

IGFBP, insulin-like growth factor binding protein; ALS, acid-labile subunit; GH, growth hormone; RACE, rapid amplification of cDNA ends; EF-1α, elongation factor-1α; Imp-L2, imaginal morphogenesis protein-late 2

1. Introduction

Insulin-like growth factor (IGF)-I is an important regulator of early development and a potent mitogen essential for normal postnatal growth in mammals. IGF-I exerts its growth-promoting action through endocrine, paracrine and autocrine mechanisms [26,31]. In mammals, IGF-I is bound to a family of six IGF-binding proteins (IGFBPs), and they control the availability of IGF-I to receptors [15,19,21,35]. Among six IGFBPs, IGFBP-3 is most abundant in the mammalian circulation and carries approximately 80% of circulating IGF-I by forming a ternary complex with an acid-labile subunit (ALS) [3,6,35]. The ternary complex prolongs the half-life of circulating IGF-I from 5 min (free form) to approximately 12 hours, and thus creates a reservoir of IGF-I. Circulating IGFBP-3 levels are generally high under positive nutritional status and up-regulated by growth hormone (GH) [3]. Human IGFBP-3 is N-glycosylated and appears as doublet bands around 40-45 kDa on electrophoresis gels [14,15]. Although binding protein glycosylation has no significant effect on the binding of IGF-I or ALS, it presumably protects IGFBP-3 from proteolysis in the circulation and affects cell surface association [14,15,30].

The IGF/IGFBP system is believed to be conserved in other vertebrates including teleosts, and multiple sequences of the IGFBP family have been identified in these groups [25,29,36,45]. In the circulation of several fish species, three IGFBP bands are typically detected by ligand blotting [25]. Their molecular weight ranges are 20-25 kDa, 29-32 kDa and 40-50 kDa. Many studies report that these fish IGFBPs in the circulation are under control of nutritional status, hormones and stress as in mammals, and suggest that physiological regulation of fish IGFBPs is also conserved [25,45]. However, it is not clear which fish IGFBPs detected by Western ligand blot corresponds to which mammalian counterparts. Identities of circulating fish IGFBPs have been assumed based mainly on the molecular weights and physiological regulation, although it is difficult to assign their specific homologies to mammalian forms of IGFBPs as these characters overlap among members of the protein family.

Fish 40-50 kDa IGFBP is a strong candidate for the mammalian IGFBP-3 ortholog [24,32,38]; however, as we will show, the major circulating 40-50 kDa IGFBP in salmon is a paralog of IGFBP-2. We previously purified 41-kDa IGFBP from serum of Chinook salmon (*Oncorhynchus tshawytscha*) and found that it is N-glycosylated as is mammalian IGFBP-3 [40]. We also reported that salmon 41-kDa IGFBP is induced by nutritional input and GH-injection

[38,39]. Moreover, circulating levels of 41-kDa IGFBP are positively well correlated with those of IGF-I and individual growth rates [4]. All of these findings suggest that the 41-kDa IGFBP in salmon is homologous to IGFBP-3 in mammals, although one conflicting observation is its partial amino acid sequence. The partial N-terminal amino acid sequence of purified 41-kDa IGFBP (20 aa) was unexpectedly most similar to IGFBP-2 [40]. Because the N-termini of IGFBPs are well conserved among six IGFBPs, we were unable to definitely conclude the identity of salmon 41-kDa IGFBP. Recently, five IGFBP cDNAs of rainbow trout (*Oncorhynchus mykiss*) were cloned and one of them had a N-terminal amino acid sequence identical to Chinook salmon 41-kDa IGFBP [22]. However, it was placed in the IGFBP-2 clade in their phylogenic analysis while its bootstrap value was relatively low (50%). The authors assigned it as "IGFBP-3" based on the molecular weight, type of glycosylation and physiological responses [22]. This conclusion was also supported by the fact that there was no sequence of IGFBP-3 found in 350,000 ESTs for salmon and trout [22]. On the other hand, Rodgers et al. [37] comprehensively analyzed the sequences of available vertebrate IGFBPs and IGFBP-related proteins and pointed out that trout "IGFBP-3" should be annotated as a paralog of IGFBP-2. Another phylogenetic analysis suggests that vertebrate IGFBPs have eight subfamilies and fish "IGFBP-3" is one of them [17]. Because there is no direct comparison between the protein sequence and cDNA sequence of the 41-kDa IGFBP in the same species and because IGFBP-3 has not been found in salmon, the identity and character of the 41-kDa IGFBP are still not clear.

In the present study, we cloned cDNA of Chinook salmon 41-kDa IGFBP and compared the sequence with the N-terminal and internal sequences of purified 41-kDa IGFBP. We also cloned cDNA of salmon IGFBP-3 for the first time. The comparison of these sequences demonstrates that salmon 41-kDa IGFBP is less like mammalian IGFBP-3 and is, in fact, a subtype of IGFBP-2. We named salmon 41-kDa IGFBP as IGFBP-2b and another paralog as IGFBP-2a. In addition, there appear to be alternative splicing forms for salmon IGFBP-2s, which provides a unique model to analyze the molecular evolution of IGFBPs in vertebrates.

2. Materials and Methods

2.1. Purification and amino acid analyses of 41-kDa IGFBP

41-kDa IGFBP was purified from serum of spawning Chinook salmon and its partial N-terminal

amino acid was determined as described in Shimizu et al. [40]. In order to analyze internal amino acids, three micrograms of purified protein were run on SDS-PAGE under reducing conditions, electroblotted onto a PVDF membrane, and stained with CBB R-250. The 41-kDa IGFBP band was excised, digest by cyanogen bromide and analyzed for amino acids of the resulting fragments at Midwest Analytical (St. Louis, MO). The amino acid mixture from each Edman degradation step was compared with the cDNA sequences of salmon IGFBPs using the FASTF algorithm [27].

2.2. cDNA cloning of salmon IGFBPs

Liver cDNA was prepared from a 2-year-old male Chinook salmon. Degenerate forward primers for salmon 41-kDa IGFBP were designed from the N-terminal amino acid sequence of purified protein (5' GTITTYTAYTGYCCIAARTGYACNGC 3', where I indicates inosine; Y = C or T; R = A or G; N = any base), and a degenerate reverse primer was designed from the C-terminal region conserved among the IGFBP family (5' TGYCCRTAYTTRTCCACRCACCAGCA 3'). RT-PCR was performed with a Perkin Elmer Gene Amp Thermal Cycler (Perkin Elmer Cetus, Foster City, CA) and components from Novagen (Madison, WI). PCR cycles consisted of 1 cycle of 94°C for 3 min; 36 cycles of 94°C for 45 sec, 50°C for 45 sec, 72°C for 1 min 30 sec; 1 cycle of 72°C for 5 min. PCR products were cloned into the pSTBlue-1 Blunt Vector using the Perfectly Blunt Cloning Kits (Novagen) and sequenced as described in Shimizu et al. [41]. Gene specific primers for 41-kDa IGFBP were designed from the sequence of the partial cDNA (Foward: 5' GTACCCAACCGCACTGAAGAGCACCGG 5'; Reverse: 5' TGGTTTTGAGCTCGTTCTGGGCCTGC 3'). Full-length cDNA was obtained by 3'- and 5'-rapid amplification of cDNA ends (RACE) using the SMART RACE cDNA amplification kit (BD Biosciences, Palo Alto, CA). A set of primers for a minor cDNA fragment (IGFBP-2a) was also designed (Forward: 5' GGAGAACGCTATGCGCCAGCACCGA 3'; Reverse: 5' TGACACTGAATCTGCTTGCCGCGGG 3') for RACE.

Heart cDNA was prepared from a spawning male Chinook salmon. The sequences of degenerate primers for salmon IGFBP-3 were originally used for the cloning of tilapia IGFBP-3 cDNA [9] with slight modifications (Forward: 5' GGTCCYGTGGTGCGCTGCGAGCC; Reverse; 5' TGYCCRTAYTTRTCCACRCACCAGCA 3'). An additional degenerate reverse primer was designed from a conserved region of IGFBP-3 for nested PCR (5'

TTDGGRATICKRAAICKICKNGGRTT 3', where D = A, G, or T; K = G or T). RT-PCR was performed with a Veriti Thermal Cycler (Applied Biosystems) and components from Promega (Madison, WI). A cDNA for zebrafish IGFBP-3 was used as a positive control. PCR products were cloned into the pGEM-T Easy Vector Systems (Promega) and positive clones were sequenced. Two sets of gene specific primers were designed for 3'-RACE (Foward 1: 5' AAACTCAACACCTTTCTGCTCCCCGCG 3'; Foward 2: 5' GGAAGGCGGAGGTGGTGGACATCGGG 3') and for 5'-RACE (Reverse 1: 5' ACCGTCTTGGTCGTCGTCACGGTGC 3'; Reverse 2: 5' GCACGGGGAGCAGAAAGGTGTTG 3'). Full-length cDNA was obtained by RACE as described above.

2.3. Analyses of salmon IGFBP sequences

Deduced amino acid sequences of salmon IGFBPs were aligned with human IGFBPs using the ClustalW method in the DNA Data Bank of Japan website (http://www.ddbj.nig.ac.jp). Signal peptide was estimated by using the SignalP 3.0 server (http://www.cbs.dtu.dk/services/SignalP) and molecular weight of mature protein was calculated by using the Compute pI/MW tool (http://us.expasy.org/tools/pi_tool.html). Potential N-glycosylation sites were detected by using the NetNGlyc 1.0 server (http://www.cbs.dtu.dk/services/NetNGlyc). The amino acid sequences of full-length IGFBPs and full-length human IGFBP-rP1 were subjected to the ClustalW analysis to create a phylogenetic tree using a neighbor-joining method (based on uncorrected p-distance). The reliability of the tree topology was assessed by the bootstrap method with 1000 replications. NJplot software was used to prepare a graphical view of the phylogenetic tree [33].

2.4. Tissue distribution of salmon IGFBPs

Various tissues are collected from male (immature and maturing) and female (immature) adult Chinook salmon. Sampling was carried out in accordance with the guidelines of the University of Washington Institutional Animal Care and Use Committee. Expression of salmon IGFBPs in various tissues was analyzed by RT-PCR. Elongation factor-1α (EF-1α) was used as a reference gene (Foward: 5' GAATACCCTCCTCTTGGTCGTTT 3'; Reverse: 5' TCGACGGCCTTGATGACA 3'). New primer pairs were designed for IGFBP-2a (Forward: 5' TAAATGACAAGACGTTCCACGC 3'; Reverse: 5' CTATTTCTGGGCTGAGACGAG 3') and

IGFBP-2b (Forward: 5' AGAATGGTGCTATATTTTAGCTGCG 3'; Reverse: 5' TTATATCTCTGCCATCTGCAGGAC 3'). PCR cycle was adjusted for each gene (30, 45 and 36 cycles for EF-1α, IGFBP-2a and IGFBP-2b, respectively). RT-PCR (36 cycles at 50°C) for salmon IGFBP-3 was first performed using a primer set (Forward: 5' ATGCGTGCTTTGTCTTACTGCGGTG 3'; Reverse: 5' CCCCCAGACACTGACTCCACCTTG 3') followed by second-round PCR (25 cycles at 60°C) using a nested primer (Forward: 5' AAACTCAACACCTTTCTGCTCCCCGCG 3').

3. Results

Degenerate RT-PCR for the 41-kDa IGFBP amplified two bands (data not shown). These bands were sequenced and confirmed as different IGFBPs. Full-length cDNAs for these IGFBPs were obtained by RACE using gene-specific primers and their deduced amino acid sequences are shown in Figure 1. One of the two IGFBPs has an N-terminal sequence identical to that of the 41-kDa IGFBP. Moreover, amino acids obtained after digesting purified 41-kDa IGFBP by cyanogen bromide were assigned to the internal regions of the cloned cDNA (Fig. 1), showing that the salmon IGFBP-2b is the cDNA for 41-kDa IGFBP. Both 41-kDa IGFBP and the other IGFBP had high sequence identity with human IGFBP-2 (Table 1) and were placed in the IGFBP-2 clade in the phylogenetic analysis (Fig. 2). Thus, they are co-orthologs of mammalian IGFBP-2 sharing 56% sequence identity. Based on the degree of sequence identity with human counterparts, the 41-kDa IGFBP was named IGFBP-2b (accession no. HM358881), and the other as IGFBP-2a (accession no. HM358880). The 41-kDa IGFBP (IGFBP-2b) had three potential N-glycosylation sites, whereas one site was found for IGFBP-2a (Fig. 1). Both salmon IGFBP-2s have the Arg-Gly-Asp (RGD) integrin recognition site (Fig. 1).

We next sought to clone cDNA for salmon IGFBP-3 from the liver based on the conserved IGFBP-3 sequences among different species. Our first attempt was unsuccessful; RT-PCR amplified IGFBP-5 (accession no. HM536184) but not IGFBP-3, probably due to the low expression of IGFBP-3. We then used a cDNA template from the heart and designed a new degenerate primer for nested PCR. After a second-round of PCR, a partial cDNA for IGFBP-3 was amplified and a full-length cDNA was obtained by 5'- and 3'-RACE (Fig. 1). A cloned cDNA (accession no. HM536183) had high sequence identity with human IGFBP-3 (Table 1), and had the

195 motifs typical for mammalian IGFBP-3 such as the basic C-terminal region for ALS and heparin
196 binding, two N-glycosylation sites and possible nuclear localization signal (Fig. 1). Its identity as
197 salmon IGFBP-3 was also confirmed by its position in the phylogenetic tree (Fig. 2).
198 Salmon IGFBP-2a and IGFBP-2b (41-kDa IGFBP) were highly expressed in the liver
199 and also detected in other tissues (Fig. 3). Although IGFBP-3 was expressed in a variety of tissues,
200 it was visible only after a second-round of PCR. The liver showed little or no expression of
201 IGFBP-3 (Fig. 3).
202 During the cloning of salmon IGFBP-2s, we found PCR products that differed in size
203 (Fig. 4). Smaller products are most likely alternative splicing forms that retain N- and C-termini
204 but lack part of the mid region (Fig. 5). However, splicing sites were different between
205 short-forms of IGFBP-2a and -2b; short IGFBP-2b (accession no. HM536182) lacks putative exon 3
206 whereas the splicing site for short IGFBP-2a (accession no. HM536181) spans parts of exon 1 and 2
207 (Fig. 5).
208
209 **4. Discussion**
210 The mammalian IGF system consists of two ligands (IGF-I and IGF-II), two receptors (type I and
211 type II) and six IGFBPs (IGFBP-1-6). All of the components appear to be conserved among
212 vertebrates including teleosts [25,36,45]. The 40-50 kDa IGFBP that is visible on Western ligand
213 blots of fish plasma has been assumed to be IGFBP-3 based on its size and response to GH and
214 fasting. In the present study, however, we demonstrate that salmon 41 kDa IGFBP is not IGFBP-3
215 but a paralog of IGFBP-2. This finding suggests that although the components of the IGF system
216 are well conserved among vertebrates, the roles that these components play in regulating growth
217 may differ among species.
218 We cloned cDNA of Chinook salmon 41-kDa IGFBP by RT-PCR using degenerate
219 primers designed from the partial N-terminal amino acid sequence of purified 41-kDa IGFBP from
220 the same species. A cloned cDNA showed the highest sequence identity with IGFBP-2, which
221 conflicts with the assumption that the 41-kDa IGFBP is the physiologic equivalent to mammalian
222 IGFBP-3, but conforms to the analysis of Rodgers et al. [37]. One concern is the possibility that
223 IGFBP-2 might be contaminated in the purified protein fraction, analyzed for N-terminal amino
224 acid sequence and cloned as 41-kDa IGFBP cDNA. This possibility is countered by the fact that

the internal amino acids of digested purified protein matched those of the cloned cDNA. Thus, 41-kDa IGFBP is indeed IGFBP-2. We found another IGFBP-2 exhibiting higher sequence identity with the human IGFBP-2 counterpart. This form and the 41-kDa IGFBP are co-orthologs of mammalian IGFBP-2, which we term IGFBP-2a and -2b, respectively. The presence of paralogs for IGFBP-2 and other IGFBP types have been reported in a wide range of teleosts including Atlantic salmon (*Salmo salar*), zebrafish (*Danio rerio*), fugu (*Takifugu rubripes* and *Tetradon nigroviridis*), stickleback (*Gasterosteus aculeatus*) and medaka (*Oryzias latipes*) [7,23,46], suggesting that gene duplication of the IGFBP family occurred before the teleost radiation. Duplication of a gene relaxes the selective pressure on its functions and is a force of molecular evolution. One (or both) of duplicated genes often undergoes nonfunctionalization, neofunctionalization or/and subfunctionalization [34]. Duan and co-workers indicated that duplicated zebrafish IGFBPs underwent temporal and spatial subfunction partitioning and proposed the utility of the zebrafish model for the study of the functions of the duplicated genes during embryonic development [12,23,44,46]. Salmon also have duplicated IGFBP genes and may be useful for analyses for postnatal growth.

Salmon 41-kDa IGFBP (IGFBP-2b) is distinct from the mammalian counterpart in a number of characteristics. First, the salmon IGFBP-2b has three N-glycosylation sites in the cDNA sequence and the mature protein is indeed N-glycosylated [40]. In humans, N-glycosylation is found only in IGFBP-3 and -4 [14,35]. The role of N-glycosylation is probably to prolong the half-life of the protein in the circulation and promote its interaction with the cell surface [14,15,30]. Addition of carbohydrates should help salmon IGFBP-2b function as a main carrier of circulating IGF-I. Second, it is up-regulated by GH and anabolic states, and is positively correlated with plasma IGF-I levels and individual growth rates as is mammalian IGFBP-3 [4,38,39]. In contrast, human IGFBP-2 is generally inhibitory to growth and increases after fasting and under several pathological conditions [5,18]. It is speculated that transcriptional regulation of 41-kDa IGFBP may be similar to those of mammalian IGFBP-3 although we have no data yet on the promoter region. An additional unique feature of salmon IGFBP-2b is the presence of a short form presumably derived from alternative splicing. The short form of IGFBP-2b lacks the portion in the mid region encoded by putative exon 3, but has a complete N-terminus and truncated C-terminus, suggesting it retains binding ability for IGFs. In addition, salmon IGFBP-2a also

possesses a short-form that lacks part of putative exon 1 and 2. These forms are detected mainly in the liver, and expression of the short forms was much lower than that of the non-spliced forms. The biological significance of the splicing variants is not known and awaits future study.

Given that salmon 41 kDa IGFBP is IGFBP-2b, a question has been whether IGFBP-3 exists in salmon. Despite searching 350,000 ESTs in the rainbow trout and Atlantic salmon databases, the sequence of IGFBP-3 could not be found [22,37]. One hypothesis was that salmon lost IGFBP-3, and one of the duplicated salmon IGFBP-2s compensate for functions similar to mammalian IGFBP-3 [22]. This hypothesis may be true with regard to the acquisition by salmon IGFBP-2b of physiological roles of IGFBP-3. However, the present study shows that salmon does have IGFBP-3. IGFBP-3 was expressed in a variety of tissues but at very low levels especially in the liver since a second PCR was necessary to visualize the band. The low level of expression might be a reason why IGFBP-3 is not found in the EST databases. The sequence of salmon IGFBP-3 is reasonably conserved including the motifs important to interact with ALS and heparin [16], but it is not a main carrier of circulating IGFs as none of the three major IGFBPs in the salmon circulation correspond to IGFBP-3 [40,41,unpublished data]. To explain this situation, the difference in the organization of livers between mammals and fishes needs to be considered. In mammalian livers, Kupffer cells and endothelial cells are the sites of IGFBP-3 production [10,43]. In contrast, Kupffer cells are rarely observed in fish livers [8], which may account for the low expression of IGFBP-3 in the liver of salmon.

One of the important roles of IGFBP-3 is to stabilize IGFs in the circulation by forming a large-molecular weight ternary complex with ALS, so that IGFs do not cross the endothelial barrier and a large pool of IGFs can be maintained [3,35]. Due to the presence of the ternary complex, circulating IGF levels are high in humans (200-300 ng/ml for IGF-I, 400-600 ng/ml for IGF-II) [3,35]. In contrast, IGF levels in non-mammalian species such as chicken and salmon are relatively low, approximately 20-50 ng/ml under normal physiological conditions [25,29,38]. Chicken has a large molecular weight complex of IGF since IGF-binding activity is found around 150 kDa on gel filtration [28], although the presence of the ternary complex consisting of IGF, IGFBP-3 and ALS has not been demonstrated. In teleosts and lamprey, there is no evidence for the ternary complex based on the molecular distribution of IGF-binding activity [13,42] and endogenous IGF-I [38]. These studies suggest that the ternary complex is not present in fishes.

285 However, the ALS gene is present and expressed in salmon and other teleosts since it is found in
286 fish genomes and EST databases (Fig. S1). The sequences of zebrafish and trout ALSs are also
287 well conserved showing 52 and 55% identity with human counterpart, respectively (Fig. S1). In
288 addition, these ALSs have five to six potential N-glycosylation sites, which are important to
289 maintain the negative charge of ALS to bind IGFBP-3 in humans [20]. These reports and
290 observations suggest that salmon have all the components (i.e. IGFs, IGFBP-3 and ALS) but do not
291 form the ternary complex. The very low expression of IGFBP-3 might be a basis for the lack of a
292 ternary complex. However, the simplest possibility for the lack of ternary complex in fish is that
293 fish ALS and IGFBPs have not evolved complimentary structural attributes to bind to each other.
294 A clear contrast is *Drosophila*. The fly has insulin-like peptides, called Dilps as well as an
295 immunoglobulin superfamily molecule distantly related to mammalian IGFBP-rP1, called Imp-L2
296 (imaginal morphogenesis protein-late 2) and an ortholog of vertebrate ALS [2]. The ternary
297 complex in this species plays roles in regulating growth, carbohydrate and fat metabolism.
298 Although it remains unclear how binding of Imp-L2 to Dilps arose as well as formation of the
299 ternary complex, the important role of the ternary complex in a wide range of animals may be a
300 result of convergent evolution, but in the vertebrate lineage, the acquisition of the ternary complex
301 with ALS might be relatively recent or fish IGFBP-3 lost the ability to form the ternary complex.
302 It needs to be clarified whether the apparent lack of the ternary complex is restricted to certain fish
303 species or it holds for all teleosts.
304 Our findings may imply that IGFBP-2 was the ancestral major IGF carrier in vertebrates.
305 Six IGFBPs are thought to be derived from a single ancestral IGFBP through three to five gene
306 duplication events [1,11,36,37]. One hypothesis is that IGFBP-2 and -5 diverged earliest followed
307 by the appearance of IGFBP-1 and -3 [11]. If this is true, in the mammalian linage IGFBP-3 might
308 take over the IGFBP-2 role. However, other phylogenetic analyses suggest different gene
309 duplication pathways [1,36,37]. Thus, the hypothesis that IGFBP-2 is the ancestral major IGF
310 carrier is too speculative at present but invites future studies on the identity and function of
311 IGFBP(s) in primitive vertebrates such as the cephalochordate and agnathan. Moreover, additional
312 studies of IGFBPs in other fish species and other vertebrates are needed to understand how these
313 proteins evolved. Salmonids may not be the best representatives of teleosts with this regard
314 because they underwent an additional tetraploid event, but useful for functional study of the

duplicated IGFBPs.

In conclusion, we demonstrate that the most abundant IGFBP in Chinook salmon serum is not a salmonid ortholog of IGFBP-3, but rather a co-ortholog of IGFBP-2, which we have termed IGFBP-2b. A second co-ortholog, termed IGFBP-2a could also be identified. The molecular expression of Chinook salmon IGFBP-3 was shown to be extremely low, and there is no evidence of a ternary complex, leading us to speculate that salmon IGFBP-3 and ALS have not evolved binding relationships. It is not known why IGFBP-2b and IGFBP-3 have different roles in salmon compared to mammals, but salmon offers a unique model to investigate the molecular evolution of IGFBPs and functional divergence of the IGF system in vertebrates.

Acknowledgments

We thank Dr. Cunming Duan, Department of Molecular, Cellular, and Developmental Biology, University of Michigan, for providing cDNA for zebrafish IGFBP-3, and Penny Swanson and Adam Luckenbach of Northwest Fisheries Science Center for comments on the manuscript. This work was supported by a grant from the US Department of Agriculture, NRICGP, Animal Growth and Nutrient Utilization Program (2003-03314); and grant-in-aids for Scientific Research from Japan Society for the Promotion of Science (19580198 and 21580214).

References

[1] A.A. Abbasi, K.-H. Grzeschik. An insight into the phylogenetic history of HOX linked gene families in vertebrates, BMC Evol. Biol. 7 (2007) 239.

[2] N. Arquier, C. Géminard, M. Bourouis, G. Jarretou, B. Honegger, A. Paix, P. Léopold, *Drosophila* ALS regulates growth and metabolism through functional interaction with insulin-like peptides, Cell Metab. 7 (2008) 333-338.

[3] R.C. Baxter, Insulin-like growth factor binding proteins in the human circulation: a review, Horm. Res. 42 (1994) 140-144.

[4] B.R. Beckman, M. Shimizu, B.A. Gadberry, K.A. Cooper, Response of the somatotropic axis of juvenile coho salmon to alterations in plane of nutrition with an analysis of the relationships among growth rate and circulating IGF-I and 41 kDa IGFBP, Gen. Comp. Endocrinol. 135 (2004) 334-344.

345 [5] W.F. Blum, N. Horn, J. Kratzsch, J.O.L. Jørgensen, A. Juul, D. Teale, K. Mohnike, M.B. Ranke,
346 Clinical studies of IGFBP-2 by radioimmunoassay, Growth Regul. 3 (1993) 100-104.
347 [6] Y.R. Boisclair, R.P. Rhoads, I. Ueki, J. Wang, G.T. Ooi, The acid-labile subunit (ALS) of the
348 150 kDa IGF-binding protein complex: an important but forgotten component of the
349 circulating IGF system, J. Endocrinol. 170 (2001) 63-70.
350 [7] N.I. Bower, X. Li, R. Taylor, I.A. Johnston, Switching to fast growth: the insulin-like growth
351 factor (IGF) system in skeletal muscle of Atlantic salmon, J. Exp. Biol. 211 (2008)
352 3859-3870.
353 [8] J. Brusle, G.I. Anadon, The structure and function of fish liver, in: J.S.D. Munshi, H.M. Dutta
354 (Eds.), Fish Morphology, Oxford & IBH Publishing Co. Pvt. Ltd., New Delhi, 1995, pp.
355 77-93.
356 [9] R. Cheng, K.-M. Chang, J.-L. Wu, Different temporal expressions of tilapia (*Oreochromis*
357 *mossambicus*) insulin-like growth factor-I and IGF binding protein-3 after growth hormone
358 induction, Mar. Biotechnol. 4 (2002) 218-225.
359 [10] E. Chin, J. Zhou, J. Dai, R.C. Baxter, C.A. Bondy, Cellular localization and regulation of gene
360 expression for components of the insulin-like growth factor ternary binding protein
361 complex, Endocrinology 134 (1994) 2498-2504.
362 [11] C. Collet, J. Candy, V. Sara, Evolutionary aspects of the IGF system, in: K. Takano, N. Hizuka,
363 S-I. Takahashi (Eds.), Molecular Mechanisms to Regulate the Activities of Insulin-like
364 Growth Factors, Elsevier Science B. V., 1998, pp. 215-223.
365 [12] W. Dai, H. Kamei, Y. Zhao, J. Ding, Z. Du, C. Duan, Duplicated zebrafish insulin-like growth
366 factor binding protein-5 genes with split functional domains: evidence for evolutionary
367 conserved IGF binding, nuclear localization, and transactivation activity, FASEB J. 24
368 (2010) 2020-2029.
369 [13] B. Degger, Z. Upton, K. Soole, C. Collet, N. Richardson, Comparison of recombinant
370 barramundi and human insulin-like growth factor (IGF)-I in juvenile barramundi (*Lates*
371 *calcarifer*): *in vivo* metabolic effects, association with circulating IGF-binding proteins, and
372 tissue localization, Gen. Comp. Endocrinol. 117 (2000) 395-403.
373 [14] S.M. Firth, R.C. Baxter, Characterisation of recombinant glycosylation variants of insulin-like
374 growth factor binding protein-3, J. Endocrinol. 160 (1999) 379-387.

375 [15] S.M. Firth, R.C. Baxter, Cellular actions of the insulin-like growth factor binding proteins,
376 Endocr. Rev. 23 (2002) 824-854.
377 [16] S.M. Firth, U. Ganeshprasad, R.C. Baxter, Structural determinants of ligand and cell surface
378 binding of insulin-like growth factor-binding protein-3, J. Biol. Chem. 273 (1998)
379 2631-2638.
380 [17] P.V. Gordon, M. Marcinkiewicz, An analysis of IGFBP evolution, Growth Horm, IGF Res. 18
381 (2008) 284-290.
382 [18] A. Hoeflich, M. Wu, S. Mohan, J. Föll, R. Wanke, T. Froehlich, G.J. Arnold, H. Lahm, H.J.
383 Kolb, E. Wolf, Overexpression of insulin-like growth factor-binding protein-2 in transgenic
384 mice reduces postnatal body weight gain, Endocrinology 140 (1999) 5488-5496.
385 [19] V. Hwa, Y. Oh, R.G. Rosenfeld, The insulin-like growth factor-binding protein (IGFBP)
386 superfamily, Endocr. Rev. 20 (1999) 761-787.
387 [20] J.B.M. Janosi, S.M. Firth, J.J. Bond, R.C. Baxter, J.D. Delhanty, N-Linked glycosylation and
388 sialylation of the acid-labile subunit. Role in complex formation with insulin-like growth
389 factor (IGF)-binding protein-3 and the IGFs, J. Biol. Chem. 274 (1999) 5292-5298.
390 [21] J.I. Jones, D.R. Clemmons, Insulin-like growth factors and their binding proteins: biological
391 actions, Endocr. Rev. 16 (1995) 3-34.
392 [22] B.B. Kamangar, J.-C. Gabillard, J. Bobe, Insulin-like growth factor-binding protein (IGFBP)-1,
393 -2, -3, -4, -5, and -6 and IGFBP-related protein 1 during rainbow trout postvitellogenesis
394 and oocyte maturation: molecular characterization, expression profiles, and hormonal
395 regulation, Endocrinology 147 (2006) 2399-2410.
396 [23] H. Kamei, L. Lu, S. Jiao, Y. Li, C. Gyrup, L.S. Laursen, C. Oxvig, J. Zhou, C. Duan,
397 Duplication and diversification of the hypoxia-inducible IGFBP-1 gene in zebrafish, PLoS
398 One 3 (2008) e3091.
399 [24] K.M. Kelley, K. Siharath, H.A. Bern, Identification of insulin-like growth factor-binding
400 proteins in the circulation of four teleost fish species, J. Exp. Zool. 263 (1992) 220-224.
401 [25] K.M. Kelley, T.D. Price, M.M. Galima, K. Sak, J.A. Reyes, O. Zepeda, R. Hagstrom, T.A.
402 Truong, C.G. Lowe, Insulin-like growth factor-binding proteins (IGFBPs) in fish: beacons
403 for (disrupted) growth endocrine physiology, in: M. Reinecke, G. Zaccone, B.G. Kapoor
404 (Eds.), Fish Endocrinology, Science Publishers, Enfield, New Hampshire, 2006, pp.

167-195.

[26] D. Le Roith, C. Bondy, S. Yakar, J.-L. Liu, A. Butler, The somatomedin hypothesis: 2001, Endocr. Rev. 22 (2001) 53-74.

[27] A.J. Mackey, T.A.J. Haystead, W.R. Pearson, Getting more from less, Mol. Cell. Proteom. 1 (2002) 139-147.

[28] J.P. McMurtry, G.L. Francis, Z. Upton, P.E. Walton, G. Rosselot, T.J. Caperna, D.M. Brocht, Plasma clearance and tissue distribution of labelled chicken and human IGF-I and IGF-II in the chicken, J. Endocrinol. 150 (1996) 149-160.

[29] J.P. McMurtry, G.L. Francis, Z. Upton, Insulin-like growth factors in poultry, Domest. Anim. Endocrinol. 14 (1997) 199-229.

[30] B.S. Miller, M.J. Khosravi, M.C. Patterson, C.A. Conover, IGF system in children with congenital disorders of glycosylation, Clin. Endocrinol. 70 (2009) 892-897.

[31] C. Ohlsson, S. Mohan, K. Sjögren, Å. Tivesten, J. Isgaard, O. Isaksson, J.-O. Jansson, J. Svensson, The role of liver-derived insulin-like growth factor-I, Endocr. Rev. 30 (2009) 494-535.

[32] R. Park, B.S. Shepherd, R.S. Nishioka, E.G. Grau, H.A. Bern, Effects of homologous pituitary hormone treatment on serum insulin-like growth-factor-binding proteins (IGFBPs) in hypophysectomized tilapia, *Oreochromis mossambicus*, with special reference to a novel 20-kDa IGFBP, Gen. Comp. Endocrinol. 117 (2000) 404-412.

[33] G. Perriere, M. Gouy, WWW-query: an on-line retrieval system for biological sequence banks, Biochimie 78 (1996) 364-369.

[34] J. Postlethwait, A. Amores, W. Cresko, A. Singer, Y.-L. Yan, Subfunction partitioning, the teleost radiation and the annotation of the human genome, Trends Genet. 20 (2004) 481-490.

[35] S. Rajaram, D.J. Baylink, S. Mohan, Insulin-like growth factor-binding proteins in serum and other biological fluids: regulation and functions, Endocr. Rev. 18 (1997) 801-831.

[36] M. Reinecke, C. Collet, The phylogeny of the insulin-like growth factors, Int. Rev. Cytol. 183 (1998)1-94.

[37] B.D. Rodgers, E.H. Roalson, C. Thompson, Phylogenetic analysis of the insulin-like growth factor binding protein (IGFBP) and IGFBP-related protein gene families, Gen. Comp.

435 Endocrinol. 155 (2008) 201-207.
436 [38] M. Shimizu, P. Swanson, W.W. Dickhoff, Free and protein-bound insulin-like growth factor-I
437 (IGF-I) and IGF-binding proteins in plasma of coho salmon, *Oncorhynchus kisutch*, Gen.
438 Comp. Endocrinol. 115 (1999) 398-405.
439 [39] M. Shimizu, A. Hara, W.W. Dickhoff, Development of an RIA for salmon 41 kDa IGF-binding
440 protein, J. Endocrinol. 178 (2003) 275-283.
441 [40] M. Shimizu, P. Swanson, A. Hara, W.W. Dickhoff, Purification of a 41-kDa insulin-like
442 growth factor binding protein from serum of chinook salmon, *Oncorhynchus tshawytscha*,
443 Gen. Comp. Endocrinol. 132 (2003) 103-111.
444 [41] M. Shimizu, J.T. Dickey, H. Fukada, W.W. Dickhoff, Salmon serum 22 kDa insulin-like
445 growth factor-binding protein (IGFBP) is IGFBP-1, J. Endocrinol. 184 (2005) 267-276.
446 [42] Z. Upton, S.J. Chan, D.F. Steiner, J.C. Wallace, F.J. Ballard, Evolution of insulin-like growth
447 factor binding proteins, Growth Regul. 3 (1993) 29-32.
448 [43] B.C. Villafuerte, B.L. Koop, C.-I. Pao, L. Gu, G.G. Birdsong, L.S. Phillips, Coculture of
449 primary rat hepatocytes and nonparenchymal cells permits expression of insulin-like growth
450 factor binding protein-3 in vitro, Endocrinology 134 (1994) 2044-2050.
451 [44] X. Wang, L. Lu, Y. Li, M. Li, C. Chen, Q. Feng, C. Zhang, C. Duan, Molecular and functional
452 characterization of two distinct IGF binding protein-6 genes in zebrafish, Am. J. Physiol.
453 Integr. Comp. Physiol. 296 (2009) R1348-R1357.
454 [45] A.W. Wood, C. Duan, H.A. Bern, Insulin-like growth factor signaling in fish, Int. Rev. Cytol.
455 243 (2005) 215-285.
456 [46] J. Zhou, W. Li, H. Kamei, C. Duan, Duplication of the IGFBP-2 gene in teleost fish: protein
457 structure and functionality conservation and gene expression divergence, PLoS One 3
458 (2008) e3926.
459

Figure captions

Fig. 1. Comparison of deduced amino-acid sequences of salmon IGFBP-2a, -2b, and -3 with those of human counterparts. Amino-acid sequences of human IGFBP-2 (NP_000588) and IGFBP-3 (NP_000589) were obtained from GenBank. They are aligned by the ClustalW method. The cysteine residues conserved in the IGFBP family are asterisked. The N-terminal amino acid sequence of purified 41-kDa IGFBP and amino acids obtained from digestion of purified protein by cyanogen bromide are circled and assigned to corresponding positions. Potential N-glycosylation sites are in solid-lined boxes. The position of Arg-Gly-Asp (RGD) integrin recognition sequence is underlined. The 18-residue basic motifs responsible for ALS binding, heparin binding and nuclear localization are in dotted-line boxes.

Fig. 2. Phylogenetic analysis of IGFBP-2 and -3 amino acid sequences. Full-length sequences of human, mouse, chicken, zebrafish, yellowtail and trout IGFBPs and human IGFBP-rP1 were analyzed by ClustalW using a neighbor-joining method based on uncorrected p-distance. The reliability of tree topology was assessed by the bootstrap method with 1000 replications. Numbers on branches are percentage of times that the two clades branched as sisters. Scale bar shows amino acid substitution per site. Trout IGFBP type in the parenthesis is re-annotated in the present study. Salmon IGFBP-2b is equivalent to 41-kDa IGFBP.

Fig. 3. Tissue distribution of salmon IGFBP expression. Various tissues were collected from male Chinook salmon and expression was analyzed by RT-PCR. A representative result from one of three individuals is shown. Note that IGFBP-3 was amplified by two rounds of PCR. EF-1α was used as a reference gene. BP-2b is equivalent to 41-kDa IGFBP.

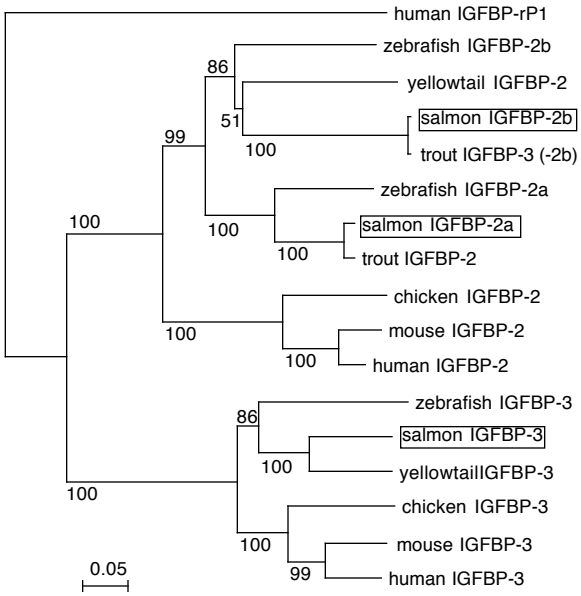
Fig. 4. Detection of short-forms of salmon IGFBP-2s. RT-PCR using liver cDNA from Chinook salmon was performed for IGFBP-2a and -2b.

Fig. 5. Deduced amino-acid sequences of splicing variants of salmon IGFBP-2s. The cysteine residues conserved in the IGFBP family are asterisked. Solid-lined boxes indicate regions missing in the splicing variants. Exon-exon boundaries indicated by vertical bars were estimated from the

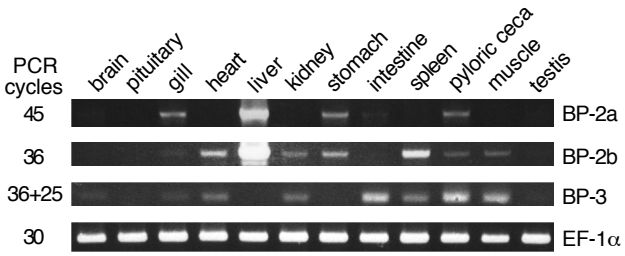
gene structures of zebrafish and human IGFBP-2.

Shimizu et al.
Figure 1

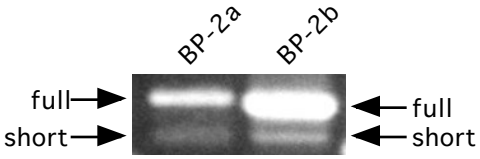
Shimizu et al.
Figure 2



Shimizu et al.
Figure 3



Shimizu et al.
Figure 4



Shimizu et al.
Figure 5

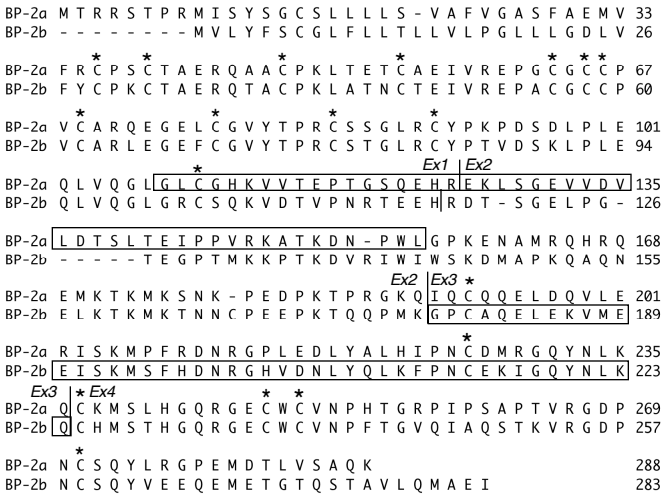


Table 1 Comparison of sequence identity and characters of salmon IGFBP-2 and -3 with human counterparts.

	Identity with hBP-2 (%)	Identity with hBP-3 (%)	Core molecular weight (kDa)	Potential N-glycosylation	NLS	RGD
human BP-2	-	32	31.4	0	-	Yes
human BP-3	32	-	28.7	3	Yes	-
salmon BP-2a	49	33	28.9	1	-	Yes
salmon BP-2b	40	30	29.2	3	-	Yes
salmon BP-3	31	52	29.5	2	Yes	-

NLS: nuclear localization signal; RGD: Arg-Gly-Asp integrin recognition sequence
Salmon BP-2b is equivalent to 41-kDa IGFBP