Evolutionary origin and divergence of the growth hormone receptor family: Insight from studies on sea lamprey

Elizabeth R. Ellens a, Jeffrey D. Kittilson a, Jeffrey A. Hall b, Stacia A. Sower b, Mark A. Sheridan a,⇑

a Department of Biological Sciences, North Dakota State University, Fargo, ND 58108, USA
b Department of Molecular, Cellular, and Biomedical Sciences, University of New Hampshire, Durham, NH 03824, USA

ABSTRACT

Sea lamprey, one of the oldest extant lineages of vertebrates, Agnatha, was used to clarify the evolutionary origin and divergence of the growth hormone receptor (GHR) family. A single full-length cDNA encoding a protein that shares amino acid identity with GHRs and prolactin receptors (PRLRs) previously characterized from teleost fish was identified. Expression of the GHR/PRLR-like transcript was widespread among tissues, including brain, pituitary, heart, liver, and skeletal muscle, which is consistent with the broad physiological roles of GH-family peptides. Phylogenetic analysis suggests that the lamprey possess an ancestral gene encoding a common GHR/PRLR that diverged to give rise to distinct GHRs and PRLRs later in the course of vertebrate evolution. After the divergence of the Actinopterygian and Sarcopterygian lineages, the GHR gene was duplicated in the Actinopterygian lineage during the fish-specific genome duplication event giving rise to two GHRs in teleosts, type 1 GHR and type 2 GHR. A single GHR gene orthologous to the teleost type 1 GHR persisted in the Sarcopterygian lineage, including the common ancestor of tetrapods. Within the teleosts, several subsequent independent duplication events occurred that led to several GHR subtypes. A revised nomenclature for vertebrate GHRs is proposed that represents the evolutionary history of the receptor family. Structural features of the receptor influence ligand binding, receptor dimerization, linkage to signal effector pathways, and, ultimately, hormone function.

© 2013 Elsevier Inc. All rights reserved.

1. Introduction

Growth hormone (GH), prolactin (PRL), and somatolactin (SL) are structurally-related hormones that have become well known for their multifunctional natures. GH, for example, regulates numerous aspects of growth, metabolism, reproduction, osmoregulation, immune function, and behavior (Forsyth and Wallis, 2002; Björnsson et al., 2004; Norrelund, 2005; Norbeck et al., 2007; Møller and Jørgensen, 2009). Approximately 300 actions have been reported for prolactin, with the most well-known being iono-osmoregulation and stimulation of milk production in mammals (Kaneko and Hirano, 1993; Wongdee and Charoenphandhu, 2012). Somatolactin, found only in teleost fish to date, appears to play a role in energy homeostasis, sexual maturation, and background color adaptation (Vega-Rubín de Celis et al., 2004; Porollo and Møller, 2007; Benedet et al., 2008; Cánepa et al., 2012). Although GH, PRL and SL are all produced by distinct cells in the pituitary, there is evidence of extra-pituitary production of all three proteins (Ben-Jonathan et al., 1996; Yang et al., 1997; Imaoka et al., 2000; Harvey, 2010). GH is a single-chain polypeptide roughly 21–22 kDa in size, and shares structural similarities with the 22–25 kDa PRL-protein and the roughly 23–24 kDa SL-protein (Rand-Weaver et al., 1992; Cavari et al., 1995; Law et al., 1996; Yang et al., 1997; Yang and Chen, 2003; Li et al., 2007; Benedet et al., 2008; Wang et al., 2010). Crystal structures of GH and PRL reveal that both have a four α-helical bundle conformation (Rowlinson et al., 2008; Agthoven et al., 2010), and although a crystal structure of SL has not been obtained, our sequence comparison suggests a structural model similar to that of GH and PRL. To date, GH but neither PRL nor SL have been detected in Agnathans, suggesting that GH is the ancestral hormone of the GH family and that the emergence of PRL and SL resulted from gene duplication events during the evolution of vertebrates (Kawauchi and Sower, 2006). The biological actions of GH, PRL, and SL result from interactions with their receptors, which also appear to be structurally related and belong to the type-1 cytokine receptor superfamily. Like all members of the this superfamily, GH receptor (GHR) family proteins are single-spanning transmembrane proteins, which possess an extracellular domain involved in receptor–receptor dimerization and hormone-binding, a single transmembrane domain, and an intracellular domain that connects the receptor to downstream

⇑Corresponding author. Address: Department of Biological Sciences P.O. Box 6050, Dept., 2715 North Dakota State University Fargo, ND 58108-6050, USA. Fax: +1 701 231 7149.
E-mail address: mark.sheridan@ndsu.edu (M.A. Sheridan).

0016-6480/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.ygcen.2013.05.008
effector pathways (Brooks et al., 2008). Based on studies in mammals, it appears that two receptors dimerize (receptor1-receptor2) to become a functional receptor complex which is then able to bind hormone (Biener et al., 2003; Brooks et al., 2008; Broutin et al., 2010). In the past decade, numerous GHR-family mRNAs have been characterized. A review of GenBank reveals some 1200 sequences for GHRs and 1700 sequences for prolactin receptors (PRLRs); although, over 50 sequences each for GHRs and PRLRs sequences for teleosts have been reported, most sequences are for tetrapods, and no information has been reported from basal vertebrate groups (i.e., Agnathans, Chondrichtians, Holostei, and basal teleosts such as Osteoglossomorpha). The existence of multiple types of GHRs and PRLRs (and in some cases subtypes) in teleosts supports the notion that the GHR family arose through a series of gene duplication events (Fukamachi and Meyer, 2007); however, considerable confusion over GHR nomenclature also has emerged, particularly with regard to the classification of a distinct SL receptor (SLR). In this study we used sea lamprey (Petromyzon marinus), one of the oldest extant lineages of vertebrates, Agnatha, to clarify the evolutionary origin and divergence of the GHR family.

2. Materials and methods

2.1. Experimental animals

Adult sea run lamprey was collected from the Cochecho River fish ladder in Dover, NH from May 22–29, 2008 during their spawning migration from the sea to fresh water. The lamprey were then transported to the Andaramous Fish and Invertebrate Research Laboratory in Durham, NH, where they were maintained in an artificial spawning channel with flow-through fresh water at ambient temperatures (9–18 °C) and natural photoperiod following the protocol reviewed and approved by the University of New Hampshire Institutional Animal Care and Use Committee.

2.2. RNA extraction

Total RNA was extracted using RNAzol RT as specified by the manufacturer (Molecular Research Center, Inc., Cincinnati, OH, USA). Isolated total RNA was re-hydrated in RNase-free water and eluted over oligo dt beads from the GenElute™ mRNA Mini-prep kit (Sigma–Aldrich, St. Louis, MO, USA) to isolate mRNA. mRNA was quantified by UV spectrophotometry (A260), diluted to 100 ng/μl, and stored at −80 °C for further analysis.

2.3. Oligonucleotide primers

Gene-specific primers used for isolation of cDNAs were designed by examining known GHR sequences using GeneTool software (BioTools, Inc., Edmonton, AB) and custom synthesized by Sigma–Genosys (The Woodlands, TX, USA). Additional primers for reverse transcription were provided in the SMARTER™ RACE cDNA Amplification Kit (BD Biosciences Clontech, Palo Alto, CA, USA). Gene-specific oligonucleotide primers used for real-time PCR of GHR were designed using ABI Primer Express® Version 2 software, based upon the sequence of our determined GHR sequence. Primers and probes were used for reverse transcription and PCR without further purification.

2.4. Isolation and characterization of GHR-like mRNAs

A three-phase approach was adopted for the isolation of a GHR-encoding cDNA using reverse-transcription (RT)-PCR and rapid amplification of cDNA ends (RACE)-PCR. Isolation of the cDNA sequence was accomplished using the SMARTER™ RACE cDNA Amplification Kit (BD Biosciences Clontech, Palo Alto, CA, USA) following the manufacturer’s protocol. In Phase I, mRNA from sea lamprey liver was reverse transcribed into cDNA containing the sequence tags necessary for SMARTER™ chemistry, and remaining cycles of PCR were carried out using a gene-specific (5′-GCCCAGG-TACGTCACTGC-3′) primer designed for 3′ RACE. After an initial denaturation at 95 °C for 3 min, a 35 cycle-PCR was performed with each cycle consisting of denaturation (95 °C for 30 s), annealing (65 °C for 30 s) and extension (72 °C for 30 s) steps; in the last cycle, the extension time was increased to 10 min. The PCR products were identified by electrophoresis on an agarose gel containing 1% of each OmniPur (EMD chemicals, Gibbstown, NJ, USA) and NuSieve GTG agarose (Bio-Wittaker Molecular Applications, Rockland, ME, USA) in 1× Tris–borate-EDTA (TBE) buffer followed by ethidium bromide staining. The PCR products were then cloned into the pGEM-T Easy Vector (Promega, Madison, WI, USA) and inserted into JM109 cells. Purified plasmids (75 fmol) were sequenced using the CEQ 2000 Dye Terminator Cycle Sequencing with Quick Start Kit (Beckman Coulter, Fullerton, CA, USA) according to the manufacturer’s protocol. Initial attempts at 5′ RACE, immediately after obtaining sequence from Phase 1 were unsuccessful; therefore, we adopted RT-PCR (Phase 2) in order to obtain more sequence, extending off the known 3′ end of the mRNA, so that new gene-specific primers could be designed for 5′ RACE.

In phase 2, after determining that the initial PCR reaction resulted in a cDNA with reasonable sequence identity with known GHR-encoding cDNAs, a unique set of primers were designed for RT-PCR (forward, 5′-CTCGCTGCTAGTACATT-3′; reverse, 5′-GCCCGGTGCTAGTACATT-3′). The forward primer was designed based on a review of coding regions of known GHR sequences together with ESTs that became available in the lamprey genome during the course of our characterization of the phase 1 product; the reverse primer for RT-PCR was designed based on the Phase 1 product. The resulting Phase 2 PCR product was visualized, cloned into the pGEM-T Easy Vector, and sequenced as described previously. After obtaining the additional sequence in phase 2, new gene-specific primers were designed and used in Phase 3, 5′ RACE. 5′-RACE was performed using a SMARTER™ RACE cDNA Amplification Kit under manufacturer’s suggested conditions with the GHR gene-specific primer (5′-GGAGACTGTGCTGGCCAGG-3′) designed for 5′ RACE. The resulting PCR products were visualized, cloned into the pGEM-T Easy Vector, and sequenced as described previously.

2.5. Quantification of GHR-encoding mRNA

2.5.1. Preparation of cDNA standards

Approximately 1 μg of known cDNA product was used as template for PCR with forward and reverse gene-specific primers (primers used in phase 2). Following an initial denaturation cycle of 94 °C for 5 min, 35 PCR cycles were performed; each consisting of 1 min denaturation (94 °C), 1 min of annealing (42 °C), and 1 min of extension (72 °C). In the last cycle, the extension time was increased to 10 min to ensure complete extension. The resulting PCR products were visualized under ultraviolet light, cloned into the pGEM-T easy vector, and their sequences verified as described previously. The cDNA standard (purified from cloned plasmid containing the cDNA) was used to generate the standard curve. With the sequence known, the cDNA (standard) was quantified by UV (A260) spectrophotometry and converted to molecule number based upon molecular weight and was then serially diluted. The average threshold cycle number (Ct) value from each dilution was plotted against the absolute amount of standard in the sample to generate the standard curve.

2.5.2. Real-time reverse transcription PCR

The previously purified mRNA (ca. 100 ng), was reverse transcribed in a 10 μl reaction using AffinityScript QPCR cDNA Synthe-
sis Kit (Stratagene, La Jolla, CA, USA) according to the manufacturer’s instructions. Reactions without reverse transcriptase (to confirm the absence of DNA contamination in the RNA) or without template (to confirm reagents were not contaminated with template from earlier experiments) were included as the negative controls; no amplification was detected in either negative control. Melting point analysis was carried out to verify amplification specificity under real-time PCR conditions. A single peak melting/dissociation curve confirmed the absence of additional DNA species such as primer–dimers.

Real-time PCR reactions were carried out for samples, standards, and controls in a 10 μl reaction; each reaction contained 1 μl cDNA, 1 μl each of forward primer, reverse primer and probe at concentrations optimized for the mRNA species to be measured, 1 μl RNase-free deionized water, and 5 μl Brilliant II SYBR® Green Low ROX QPCR Master Mix (Agilent Technologies®, Santa Clara, CA, USA) The primers used (forward, 5’-TGCAGCGAAGTTTCTTCTTCT-3‘; reverse, 5’-GGAGGTGCAATGTCATC-3’) were designed to be gene specific but would amplify both truncated GHR/PRL and the splice variant, thus detecting total lamprey GHR/PRLR expression. Cycling parameters for real-time PCR were as follows: 95 °C for 10 min, and 45 cycles for 95 °C for 30 s plus 59 °C for 1 min.

2.5.3. Data analysis

Copy number calculations were based on threshold cycle number (CT). The CT for each sample was determined by the MX3000P™ real time analysis detection software after manually setting the threshold. Sample mRNA expression (copy number) was determined by relating CT to the standard curve (the preparation of which was described above). No template control samples did not exceed a maximal increase of 300 (ΔRn) fluorescence units over 45 cycles. Therefore, copy numbers of mRNA were considered non-significant if CT exceeded 45 cycles; this value corresponds to a detection limit of less than 100 mRNA copies.

Quantitative data are expressed as means ± SEM. Statistical differences were estimated by a cell means model ANOVA followed by Duncan’s multiple range test; a probability level of 0.05 was used to indicate significance. All statistics were performed using Sigma Stat v. 1.0 (SPSS, Chicago, IL, USA).

2.6. Sequence analyses and modeling

The nucleotide and associated protein sequences were aligned and analyzed with GeneTool and PepTool sequence analysis programs, respectively (BioTools Inc., Edmonton, Alberta, Canada). Clustal X (default parameters, except corrected for multiple substitutions) was used in conjunction with the neighbor-joining method to generate the phylogenetic tree; the tree was visualized with TreeView and rooted to the rat erythropoietin receptor. Only completely overlapping sequences over 300 nt were used in the analysis. Potential phosphorylation sites (either protein kinase A or protein kinase C) were predicted by NetPhos 2.0 (default settings). Synteny maps were constructed by manually annotating conserved genes flanking GHR loci based on available databases.

To produce a 3-dimensional model of the lamprey protein, deduced from the obtained mRNA sequence, SWISS–MODEL (a protein structure homology-modeling server workspace; Bordoli and Schwede, 2012) was used to identify a suitable template structure by comparison of the amino acid sequence of interest with Protein Data Bank entries and to subsequently generate a model. Human GHR (PDB ID: 3hhr chains B and C) was used as the template structure for the homology modeling of the extracellular domain of the lamprey GHR/PRLR as well as for the trout GHRs. Rat PRLR (PDB ID: 3npz chains B and C) was used as the template protein for the trout PRLR target model.

3. Results and discussion

3.1. Characterization and distribution of GHR/PRLR mRNA

A cDNA fragment 1115 bp in length, including the poly-A tail, was characterized in Phase 1 (3’ RACE). Sequence analysis suggested the successful isolation of a fragment of a GHR/PRLR-like protein. In phase 2 we designed a forward primer based on known GHRs together with ESTs that became available in the lamprey genome and a reverse primer based on the sequence of the phase 1 PCR product to conduct RT-PCR (phase 2). The resulting cDNA obtained was 365 bp in length and overlapped with the cDNA fragment obtained in phase 1. The final phase, phase 3, of our
approach consisted of 5′RACE PCR, using a gene-specific primer designed from the sequence of the phase 2 PCR fragment. Fig. 1 illustrates the 3-phase approach and the assembly of the resulting fragments.

Assembly of the phase 1, phase 2 and 508-bp phase 3 PCR fragments yielded a unique full-length 1711-bp cDNA (Fig. 1d). The sequence contains an initiation site 211-bp from the most 5′ end, a 648-bp coding region, and a 849-bp 3′ untranslated region with a poly-A tail (Fig. 2). The coding region encodes a 216-amino acid protein containing the entire extracellular domain of a GHR/PRLR-like protein, including the hormone binding surface; however, both the transmembrane domain and intracellular domain are absent, explaining a shorter mRNA length than predicted (GHR mRNAs are typically about 2500–3000 bp in length, in most

Fig. 2. The cDNA and deduced amino acid sequences of truncated GHR/PRLR (seq 1) and a GHR/PRLR 5′ splice variant (seq 2) isolated from sea lamprey. The predicted hormone-binding regions are boxed in color (region 1 = magenta, region 2 = cyan, region 3 = green, region 4 = blue); the predicted dimerization domain is boxed in gray; and possible phosphorylation sites are denoted by ●. The position of putative splice sites is indicated by □.
Therefore, we have designated this segment truncated lamprey GHR/PRLR. The truncated lamprey GHR/PRLR mRNA shares (using only overlapping segments in comparison) 35–36% nucleotide identity to trout GHRs and 37% nucleotide identity to trout PRLR. We also obtained a second 5’ RACE 473-bp product; this product shows two deletions and one insertion compared to the first 5’-RACE product and does not overlap with the products from phase 1 or phase 2 (Fig. 1e, Fig. 2). With the recent update of the lamprey genome database following the characterization of all of our cDNAs, we found that this second 5’-RACE fragment is part of the same contig as the other fragments isolated, thus supporting the notion that this second fragment is derived from alternative splicing (putative splice sites denoted on Fig. 2); therefore, we have designated this fragment as the lamprey GHR/PRLR 5’ splice variant. Unfortunately, high G–C content prevented us from obtaining the balance of the sequence of this variant. Given that exhaustive screening of hundreds of colonies from dozens of PCRs yielded only these sequences, we surmise that the splice variant may encode the full-length receptor with the transmembrane and intracellular domains.

Truncated forms of GHR are derived by several mechanisms and the resulting soluble protein may serve numerous roles, including as GH binding proteins (GHBP) that affect the distribution and bioavailability of GH. The first-discovered GHBP were found to be produced through proteolytic cleavage of the extracellular domain from the membrane-bound receptor (Leung et al., 1987; Baumann, 2001). Later, it was found that GHBP also may derive through alternative splicing of the gene encoding the full-length GHR (Edens et al., 1994). Such alternative splicing of the full-length GHR gene was reported in fish and mammals, including in turbot (Calduch-Giner et al., 2001), Chinese sturgeon (Acipenser sinensis) (Liao and Zhu, 2004), and rodents (Baumbach et al., 1989; Smith et al., 1989; Edens et al., 1994; Martini et al., 1997; Edens and Talamantes, 1998; Talamantes and Ortiz, 2002; Liao and Zhu, 2004).

Interestingly, alternative splicing of the GHR gene has been found to produce distinct mRNAs that encode other “truncated receptors” besides the soluble GHBP. In humans, alternative splicing of the GHR can produce a membrane bound receptor lacking the entire intracellular domain that is seen in the full-length receptor, and instead exhibits 6 intracellular amino acids that are novel, after the transmembrane domain and before the stop codon (Ross et al., 1997). A second truncated hGHR was found to be missing 26 amino acids of exon 9 in which case a frameshift encoding an alternative amino acid sequence results in a premature stop codon, thus

Fig. 3. Distribution of GHR/PRLR mRNA among tissues in adult sea lamprey. Expression of mRNA was evaluated by real-time PCR. (A) Qualitative expression was assessed by separating PCR amplicons (after 45 PCR cycles) on agarose visualized by ethidium bromide staining and transillumination. mRNA expression was quantified by determining the threshold cycle number (CT) for each sample from its amplification plot, as exemplified for liver (B). Sample copy number was then determined by relating CT to a standard curve (C) comprised of a serial dilution of a known amount of lamprey GHR/PRL cDNA, then normalized to β-actin and (D) expressed as mean ± SEM (n = 4–6). Groups with different letters are significantly different (p < 0.05).
truncating the majority of the intracellular domain (Dasot et al., 1996; Ross et al., 1997). Besides the soluble GHBP produced by alternative splicing, it appears that most of the truncated receptors that have been characterized result in a receptor with an altered intracellular domain. However, in several human prostate cancer cell lines, an expressed truncated receptor was found to be missing exon 3, which encodes the N-terminal portion of the extracellular domain to result in a receptor that is 22 amino acids shorter in length and that has an unknown physiological relevance (Chopin et al., 2002). Alternative splicing has also been found to be responsible for a long form of GHR in black sea bream, which includes an additional alternative splice site not characterized in any other organism (Tse et al., 2003). This long form of GHR includes an additional 33-amino acid segment in the intracellular domain that does not result in a frameshift to cause any other differences in the intracellular domain. Alternative splicing of the PRLR gene has also been found to produce multiple transcripts, including both truncated and long forms of membrane bound PRLRs (Tanaka et al., 2000).

Expression of the lamprey GHR/PRLR mRNA among tissues is shown in Fig. 3. Lamprey GHR/PRLR mRNA is widely expressed among tissues; it was detected in brain, pituitary, thyroid, eye skin, intestine, liver, kidney, heart, skeletal muscle, ovary, and testis (Fig. 3a). Such wide distribution is consistent with the pleiotropic actions of GH family peptides (Kaneko and Hirano, 1993; Forsyth and Wallis, 2002; Björnsson et al., 2004; Vega-Rubín de Celis et al., 2004; Norrelund, 2005; Norbeck et al., 2007; Porollo and Meller, 2007; Benedet et al., 2008; Moller and Jørgensen, 2009; Cáñepa et al., 2012; Wongdee and Charoenphandhu, 2012). The highest expression of lamprey GHR/PRLR mRNA was observed in the pituitary and liver (Fig. 3c), a finding that is similar to the expression profile of GHRs generally in fish [cf. 35, 40]. Notably, the liver also is the chief source of truncated GHRs such as GHBP in other species (Carlsson et al., 1990; Ross et al., 1997).

3.2. Evolution of GHRs

Fig. 4 depicts the phylogenetic relationships among the GHR/PRLR family, with emphasis on GHRs in fish. The pattern that emerges is that the GHR family arose through a series of gene duplication events. The GHR/PRLR of lamprey appears to be an ancestral receptor of both GHRs and PRLRs in vertebrates. Although Chondrichthians and a common ancestral for the Actinopterygian and Sarcopterygian lineages diverged from early Gnathostome stock, the pattern of GHR/PRLR divergence is not clear because of the absence of sequence information in any Chondrichthian representative. (Although work on the elephant shark genome has begun, it is incomplete, and blasts with known GHRs or PRLRs results in no hits.) What is known is that after the divergence from PRLR, the common ancestral GHR for the Actinopterygian and Sarcopterygian lineages subsequently diverged. Descendants of the early Sarcopterygians retained a single copy of the GHR gene as seen in lungfish and the tetrapods. The Actinopterygian lineage underwent a subsequent duplication, the fish-specific genome duplication (FGSD) or 3R event. The current hypothesis regarding the timing of the FGSD event is suggested to be sometime after the divergence of the Acipenseriformes (represented by sturgeon in Fig. 4) and the Holosteans (considered a monophyletic group containing Semionotiformes and Amiiformes) from the lineage leading to teleost fish, but sometime before the divergence of Osteoglossiformes, the most basal teleost group (Fukamachi and Meyer, 2007; Ellens and Sheridan, 2013). However, there is no sequence information for GHRs in species representing Semionotiformes (e.g., gar), Amiiformes (e.g., bowfin), or Osteoglossiformes (e.g., arawana). What is clear, is that following the FGSD event there are two distinct clades of GHRs within the teleosts, representing two distinct GHR subtypes: type 1 GHR (GHR1) and type 2 GHR (GHR2) (Fig. 4). The phylogenetic analysis also supports the notion that type GHR2 is a tetrapod-specific paralog of GHR1 and that GHR1 is the homolog of the ancestral GHR that persisted in early Actinopterygians (sturgeon) and in the Sarcopterygian lineage (lungfish, tetrapods) as suggested by Fukamachi and Meyer (2007). Several lineages of teleosts independently underwent a more recent duplication (4R) event, including the salmonids. It appears that many species of salmonoids have retained both copies of the duplicated GHR2 gene (i.e., GHR2a and GHR2b have both been characterized in rainbow trout, coho salmon, and Atlantic salmon). To date, only the Jian carp has had two GHR1s characterized, GHR1a and GHR1b [cf. 19]. Interestingly, this variety of carp was artificially developed through integrated genetic breeding techniques, using two subspecies of common carp (using Hebao red common carp, Cyprinus carpio var. Wuyuanensis, as the original maternal parent and Yuenjiang common carp, Cyprinus carpio Yuenjiang, as the original paternal parent). The common carp (Cyprinus carpio) is one of two cyprinid species that is believed to have gone through a tetraploidization event, likely independent of the salmonoids (Ellens and Sheridan, 2013). Presumably, one of the copies that resulted from the 4R duplication event in salmonoids either has been lost or has yet to be characterized.

To further explore the relationships among GHRs, a qualitative microsynteny survey was performed (Fig. 5). While the complete assembly of sequence and chromosome mapping has yet to be completed in most species of fish, we made use of all available databases to manually align genes around the GHR loci. (Unfortunately, there is insufficient information in the lamprey genome to include in the analysis.) Several patterns emerge from the synteny analysis. The descent of GHRs in fish and Sarcopterygians from a common ancestor is supported by the juxtaposition of C7 near the GHR locus (except Coelacanth, the genome of which is incomplete and data on the location of C7 is not available). The duplication and divergence of teleost GHRs also is reflected in their syntenic maps. Multiple genes retain their position near the GHR locus, including PLCXD3, C5orf51, FBX04, SEPP1, and ZNF131, in humans as compared to the Type 1 GHR genes of teleosts; however, other genes near the locus of GHR1 in teleosts, such as IP011 and OXTCT2, were moved to other locations in humans. The transposition of C6 and C7 in humans also is interesting to note. Genes such as IP5SK1B and TJP2 are only near the type 2 GHR locus in teleosts and were moved to other locations in humans. Interestingly, the CDDC152 is positioned near the Type 2 GHR locus of teleosts and retained its juxtaposition near the GHR locus in humans. The general syntenic pattern supports the notion that teleost GHR1 and GHR2 are paralogs that arose through gene duplication, and that teleost GHR1 and GHR2 are orthologs to Sarcopterygian GHR (as evidenced by humans), a conclusion that is consistent with our phylogenetic analysis and with the previous analysis by Fukamachi and Meyer (2007). Unfortunately, there is insufficient information at present regarding gene relationships following subsequent duplication events in teleosts (4R) such as salmonid tetraploidization (e.g., salmonids).

3.3. Analysis of GHR and PRLR structures

Several structural elements are important for conformation and functionality of the GHR-family proteins: the number of cysteine residues, the dimerization domain, and the hormone-binding domain. Fig. 6 shows an alignment comparing the number of conserved Cys residues as well as the homologous regions of the dimerization domain from known PRLRs and GHRs in fish and selected tetrapods. The pattern emerging is that the GHR/PRL ancestrally genes possessed 7 conserved Cys residues, a conclusion supported by the lamprey truncated GHR/PRLR (although no se-
Fig. 4. Phylogenetic tree of the known growth hormone receptors (GHR) of fish and selected other vertebrates. Prolactin receptors (PRLR) from selected teleosts are included for comparison. The tree was based on the alignment of amino acid sequences using the N–J bootstrap method in Clustal X and considered only completely overlapping segments greater than 300 nt in length. The tree was rooted using the erythropoietin receptor as an out group and was visualized with TreeView. The branch lengths represent amino acid substitutions per site from a common ancestor and are proportional to the estimated time since divergence occurred. The nomenclature for a particular receptor reflects that given by the authors originally or that which appears in databases; if the sequence was not annotated or the receptor type/subtype was not specified, the designation on the tree is ours and was chosen for consistency with the phylogenetic analysis and our proposed nomenclature. We recommend abandonment of the term somatolactin receptor (SLR), and the use of letters to designate GHR subtypes in teleosts. Sequences were obtained from either GenBank (accession numbers in parentheses) or Ensembl (protein ID numbers in parentheses) as follows: Atlantic halibut GHR (DQ062814), Atlantic salmon GHR1 (NM001123576), Atlantic salmon GHR2 (NM001123594), Atlantic salmon SLR (NM001141617), black seabream GHR1 (AF622334), black seabream GHR2 (AY662234), Catla GHR (AY091178), Channel catfish GHR (DQ101502), chicken GHR (NM_001001293), Chilean flounder GHR1 (EU004149), Coelacanth GHR (ENSLACG0000000554), coho salmon GHR1 (AF403539), coho salmon GHR2 (AF403540), common carp GHR (AY741100), common carp PRLR (AY044448), frog GHR (AF193799), gilthead seabream GHR1 (AF438176), gilthead seabream GHR2 (AY573601), goldfish GHR (AF929347), goldfish PRLR (AF144012), grass carp GHR (AY283778), Japanese crucian carp GHR (ADZ13485), Japanese eel GHR1 (AB180476), Japanese eel GHR2 (AB180477), Japanese medaka SLR (NP_001098560), Jian carp GHR1 (ADC35573), Jian carp GHR2 (ADC35574), Jian carp GHR2a (ADC35575), lamprey GHR/PRLR (this sequence), lungfish GHR (EF158856), masu salmon GHR (AB071216), masu salmon SLR (AB121047), Mozambique tilapia GHR1 (AB115179), Mozambique tilapia GHR2 (EF452496), Mozambique tilapia PRLR (EU999785), Mirgal carp GHR (AY691179), Nile tilapia GHR1 (AY973232), Nile tilapia GHR2 (AY973233), Nile tilapia PRLR (L34783), opossum GHR (NM001032976), orange spotted grouper GHR1 (EF052273), orange spotted grouper GHR2 (EF052274), orangefin labeo GHR (EU147276), pigeon GHR (D84398), rainbow trout GHR1 (JQ408978), rainbow trout GHR2s (NM001124535), rainbow trout GHR2b (NM001124731), rainbow trout PRLR (AF229197), rat erythropoietin receptor (AAH89510), rat GHR (NM0017094), Rohu labeo GHR (AY693177), South American cichlid SLR (FJ208934), Southern catfish GHR2 (AY36104), Southern catfish GHR2 (AY973231), stickleback GHR (ENSGACT00000023732), sturgeon GHR (EF158851), Takifugu GHR1 (BA68396), Takifugu GHR2 (BA68397), Tetraodon GHR (ENSTNIP0000004152), tongue sole GHR1 (FJ008664), turbot GHR (AF352396), turtle GHR (AF211173), wami tilapia GHR1 (EF371466), wami tilapia GHR2 (EF371467), Wuchang bream GHRa (AF38427), Wuchang bream GHRb (AF38428), yellowfin seabream GHR2 (AEW92012), zebrafish GHRa (EU649774), zebrafish GHRb (EU649775).
Type 1 GHRs retain the ancestral condition of 7 Cys residues with some slight variability at the last position where there has been a substitution, a derived condition that also occurred in mammals. Type 2 GHRs have between 4 and 5 Cys residues, with no type 2 GHR having a Cys in the 5th or 6th position and variability in the last position as with type 1 GHRs. By contrast, all fish and tetrapod PRLRs possess only 5 conserved Cys residues, with none in the 5th or 6th positions like type 2 GHRs and with a Cys always in the last position.

The dimerization domain is the region of interaction between the two receptors of the dimer and is also important in the rotation of the receptor-pair, relative to each other, that occurs during receptor activation (Agthoven et al., 2010). It is thought that the conformational rotation that occurs to activate the receptor results in downstream signal transduction by shifting the alignment of the "linker region," the short region found in the extracellular domain just before the transmembrane domain, which subsequently realigns the transmembrane domain and thus alters the alignment and conformation of the intracellular domain (Rowlinson et al., 2008). Mutations in the linker region support this theory, with altered ratio of STAT5 to ERK1/2 signaling (Rowlinson et al., 2008).

Recently, the characterization of the dimerization interface, also known as the stem-stem interface (or site 3), of the PRLR has been completed, and upon comparison with the GHR, even though the general region of the dimerization domain is common between the two, the actual interface is different between PRLR and GHR (Agthoven et al., 2010).
Indeed, clear patterns are seen between the receptor types (clades) as well as among closely related species. In the GHR complex, most of the interacting residues are part of the initial Leu-Gln172 region (human GHR as the reference), while most of the interacting residues in the PRLR interface are part of the latter domain region, which in human PRLR stretches from Lys-His197 (Fig. 6). Although the characterized regions of the crystal structures for GHR and PRLR used as the reference are from human (GHR) and

![Diagram](image-url)
rat (PRLR) frame, with no structural reference frame (crystal structure or NMR) available from a species of fish, conservation of several residues found to be interacting residues in the human and rat receptor interfaces, is seen. For example, Asn161, Ser163 and Asp170 (hGHR) are involved in hydrogen bonding at the hGHR dimerization interface and these residues are highly conserved among the GHRs (Fig. 6). In particular, Asp is completely conserved in all GHRs; Asn and Ser are highly conserved in the type-1 GHR clade, whereas some substitutions are seen in the type-2 clade. Notably, these Ser and Asp residues also are present in the characterized lamprey sequence. Additional examination of the initial stretch that is highly involved in the GHR-GHR interface, when comparing the equivalent area for PRLR, the characterized rPRLR is two residues shorter and is therefore unable to participate in the PRLR-PRLR interaction. This absence at the interface allows the two PRLR subdomain 2 s to get closer than the two GHR subdomains (Rowlinson et al., 2008). The difference in the relative location of the receptors within a dimer pair can be seen in Fig. 7. Although it is not possible to completely predict the dimerization interfaces in these receptors in fish without a crystal structure, the presence of a residue gap in the PRLRs and not in the GHRs of these species is a conserved feature and suggests that, like in the hGHR, this stretch of the dimerization domain is also, likely, highly involved in the actual GHR-GHR interactions. What is interesting about this

Fig. 7. Three-dimensional models of the extracellular domains of the truncated growth hormone receptor (GHR)/prolactin receptor (PRLR) from sea lamprey (center) and of the GHRs and PRLR from rainbow trout. The homology models were based on deduced protein sequences using the SWISS-MODEL workspace (Bordoli and Schwede, 2012) to produce coordinates based on templates; GHR models used the crystal structure of human GH-GHR2 (PDB ID: 3hhr) as template and the PRL model used the crystal structure of human PRL-PRLR2 (PDB ID: 3ew3) as template; images were generated with POLYVIEW-3D (Porollo and Meller, 2007).
A two-residue gap is present, however, the sequence Asn–Val–Ser (Asn and Ser being residue involved in hydrogen bonding at the GHR-GHR interface) that is present in hGHR, but not PRLR, is present in the lamprey sequence.

Another region that can be examined at the molecular level to reveal trends generally consistent with the phylogenetic analysis is the ligand-binding interface, also termed site 1 (site of binding between the first receptor and hormone) and site 2 (site of binding between the second receptor and hormone). This region is improp-
tant considering the current model of GHR activation, which includes the high affinity binding of GH to the first of the receptors of the GHR dimer, initially, which is followed by the low-affinity binding of GH to the second GHR of the dimer (Brooks et al., 2008). A mutational analysis of GH using domain swapping of goldfish (Carassius auratus) GH (gfGH) and goldfish prolactin (gfPRL) also supports a model for a single GH with two binding sites bound to a dimerized receptor. The analysis and mutational study of goldfish GH supports a model whereby three discontinuous regions on the first helix and the region in the middle of the fourth helix comprises binding site 1; the third helix appears to be associated with binding site 2 (Chan et al., 2007). While it is believed that binding site 1 of the hormone is involved in the initial binding of the first GHR (GHR1), in which case the hormone has a conformational change while GHR does not have a major conformational change, site 2 of the hormone could be considered more as the functional site that binds the second GHR (GHR2) and determines the degree of the rotation of the receptors, relative to each other (Chan et al., 2007; Brooks et al., 2008; Brouin et al., 2010); GHR2 goes through a conformational change to, essentially, accommodate the binding of GH. Thus, since this binding induces this conformational change (rotation of receptors, relative to each other) that results in signaling, differing degrees of conformational change which could happen if different ligands are binding (e.g., SL binds GHR versus GH binding GHR) may likely result in differential signaling in the cell (cf. Reindl et al., 2009).

The structural characteristics of the dimerization domain that allows for the rotation of the receptors involved in the dimer, relative to each other, has been discussed; however, the area involved in triggering that rotation is critical to these receptors’ biological activity as well. In general, the ligand binds both receptors of the dimer in the same general areas of the receptor; these areas, region 1–4, are shown in Fig. 7. Focusing on binding site 2, considering its functional importance, the actual interface between the hormone and receptor involves the N-terminus of the hormone and the Gly cavity of the hormone, which is essentially a pocket with Gly as the bottom and large amino acids surrounding to create the walls of the pocket, and Trp122 of the receptor with the Trp essentially filling the pocket. Fig. 8 shows strict conservation of this Trp in all species. Comparisons of the GHR and PRLR by Brouin et al. (2010) indicated that the walls of the cavity involve residues of α-helices 1 and 3 that form a hydrogen bond network with residues from the receptor and that these residues in the receptors (GHR or PRLR) are topologically similar, but not strictly conserved, and not all interacting residues are equivalent in the alignment. For example, Ser142 (Asp120 in PRLR) and Thr142 (Thr122 in PRLR) are residues involved in this network that are in an equivalent position in the alignment. However, Trp122 (GH) is involved in hydrogen bonding, while at a different location in the alignment Gln of PRLR interacts with the same residue of the hormone.

The N-terminus of the hormone and the residue of the receptor with which it interacts is the other section that makes up site 2. Although the N-terminus of both GH and PRL are both involved in binding GH, the amino acids, which are responsible for binding, are different between the GH–GH and PRL–PRLR structures, Asp120 is suspected to be involved in the interaction in PRLR. Pro124, Arg89, Cys140, and Ser142 are believed to be involved in binding in GH. While some of these interactions show conservation (Fig. 8), there is a great deal of difference, in general, when examining regions 1–4. For example, Cys140, which is within van der Waals’ distance (Rowlinson et al., 2008) to Ile of hGH, is strictly conserved in the type-1 GHRs but is not seen in any of the type-2 GHRs, just as it is not seen in the PRLRs. Indeed, a complete picture of binding cannot be seen unless considering the molecular makeup of the hormones as well, but as evidence indicates, the actual residues involved in the interactions that result in binding may be variable (Rowlinson et al., 2008). Interestingly, site-directed mutagenesis at the functional binding site 2 of GH creates a GH antagonist with little to no biological activity, and when recombinantly engineered into a homodimer it does not only maintain the ability to bind, but does so through the use of two site 1’s, one from each of the two GH antagonists involved in the homodimer (Langenheim et al., 2006). Furthermore, this homodimer, created using two antagonists, did not share in the antagonistic effects of the monomers, but instead acted as an agonist (Langenheim et al., 2006). With the ability of GH, PRL, and SL to interchangeably bind with the GHR homologues as well as with PRLR homologues, the most notable characteristic of the hormone-binding interface on these receptors, then, is not necessarily any particular residue, but is the global ability to be versatile in its binding. The hormone binding regions of the lamprey GHR/PRLR are shown in Fig. 7. With both receptors in the dimer in a position to examine the binding domain, the differences between the receptor that binds the hormone first and the receptor that binds the hormone second becomes apparent, supporting the notion that these receptors are plastic in their ability to bind. Returning to examine the alignment (Fig. 8), some of the most conserved residues in region 1–4 are not residues that were found to be involved in interactions with the hormones, they are residues such as Ala and Val that are intermittently dispersed in these regions (e.g., Val142/Val121, Val146/Val125 in GHR/PRLR, respectively) which would be necessary to allow various residues that could potentially interact with differing ligands to be free of steric hindrance. Even in region 1, where Val is not seen in hGHR, numerous Val residues are intermittently present in this region of GHRs in fish species. In summary, both receptors in the dimer bind the hormone with the same regions (regions 1–4), but because the binding sites on the hormone are located asymmetrically, a major conformational change occurs in the receptor, not upon hormone-binding with the initial receptor of the dimer, but upon binding with the second as the receptor basically “reaches” to where it can interact with site 2 of the hormone (Fig. 7); it is this conformational change in the receptor that allows for downstream signaling effects to occur and differential degrees of conformational change results in differential signaling (Rowlinson et al., 2008).

4. Conclusions

We have characterized a full-length cDNA from sea lamprey, Petromyzon marinus, which encodes a common GHR/PRLR. That this lamprey GHR/PRLR is a precursor to Gnathostome GHRs and PRLRs is supported by several lines of evidence. First, degenerative primers designed from conserved regions of known GHRs and PRLRs yielded only the herein reported sequences after exhaustive screening. Second, the obtained lamprey sequences share similarity–nucleotide, amino acid, and overall predicted conformation (cf. Fig. 7)—to both GHRs and PRLRs. Lastly, blasting the lamprey genome with the obtained sequence yielded only one hit on the contig containing our sequences. In addition, blasting the lamprey genome with known teleost GHRs and PRLRs (e.g., rainbow trout) also yielded 1 hit: the same sequence on the contig as that hit when blasted with the herein obtained lamprey GHR/PRLR. Moreover, a blanket blast of GenBank with the obtained lamprey GHR/PRLR hits PRLRs and GHRs in all Gnathostome lineages. It should be noted that the functionality of the lamprey GHR/PRLR is not known and detailed binding studies of the expressed GHR/PRLR remain to be conducted.

4.1. Proposed evolution and divergence of GHRs

The evolutionary history of the GHR family was examined through phylogenetic and microsynteny analysis. Fig. 9 shows...
our proposed model for the divergence of GHRs and PRLRs based on current information. The timing of early whole-genome duplication events (1R and 2R) has been debated, but recent analysis places both the 1R and 2R events prior to the divergence of Agnathans and Gnathostomes (Smith et al., 2013); therefore, as has been suggested for GH (Smith et al., 2013), we speculate that only one GHR survived after the 2R duplication. In our model, PRLRs diverged early from GHRs in Gnathostomes, followed by subsequent divergence of GHRs in the Actinopterygian lineage that gave rise to teleosts, and the Sarcopterygian lineage that gave rise to tetrapods. In the absence of sequence information in any Chondrichtian, the divergence of Chondrichtian GHR (or PRLR) and the common Actinopterygian/Sarcopterygian ancestor GHR (or PRLR) from the early Gnathostome stock is unclear. The structural heterogeneity of GHRs in teleosts results from the existence of multiple genes that arose through a series of independent gene duplication events during the course of their evolution, as well as alternative transcripts of a single gene, and the alternative splicing of a single gene appears to be conserved in this ancient lineage. The two distinct GHR genes of teleosts are paralogous and form two distinct clades: Type 1 GHRs and Type 2 GHRs.

4.2. Proposed revision of GHR nomenclature

In the past several years, the question of the existence of a distinct somatolactin receptor (SLR) has emerged. In an already-unstructured nomenclature system, the system became more complex following the characterization of what appeared to be a distinct SLR from masu salmon, based on 125I-SL binding (Forsyth and Wallis, 2002) that fell within the clade with type 1 GHRs. Fukamachi and Meyer (2007) suggested that all of the teleost type GHR1s should be referred to as SLRs, and that the other major clade (type 2 GHRs, which includes what were formerly referred to as GHR1 and GHR2 of salmonids) be referred to as GHR2s. The two type 2 GHRs of salmonids most likely arose during the more recent tetraploidization (4R) event associated with the evolution of this group. Reindl et al. (2009) suggested that the binding characteristics observed in masu salmon (Fukada et al., 2005) may be a derived trait and that it may be premature to assign the label of "SLR" to all type 1 GHRs. Indeed, as will be discussed below, several type 1 GHRs retain GH binding characteristics.

There are at least two reasons for not using the term SLR to describe GHR1, especially in species other than salmon. First, the binding of GH to eel GHR1 could not be displaced by SL (Ozaki et al., 2006). Second, GH, but not SL, activated both sea bream GHR1 and GHR2 transcription reporter systems (Jiao et al., 2006). In the end, it appears that the ability for these ligands and receptors to cross-bind (e.g., GH with PRLR or PRL with GHR) and elicit cellular effects in response to cross-binding may be species specific; thus, a nomenclature scheme based of the evolutionary relationships of these receptors would be more consistent.

Given the confused state of GHR nomenclature, it is clear that a community-wide movement toward a simplified nomenclature system that better represents the evolutionary history of this receptor family is needed. To this end, we suggest a change to a system similar to that already adopted for instances of multiple genes. This system utilizes different numbers to designate genes derived from one duplication event, then different letters to designate paralogs derived from a subsequent round of duplication. In application to GHRs, such a system would use numbers to designate the different GHR types that arose in the Actinopterygian lineage (associated with 3R or FSGD); hence, in the teleosts there would be GHR1s (we urge abandonment of the term SLR to avoid confusion) and GHR2s. The addition of different letters would be added to distinguish paralogs associated with 4R duplication events (e.g., salmonids). Because the 4R event occurred independently in (and not in all) teleost lineages, we suggest coupling the species name to the gene designation. This will necessitate changes to existing names (and some temporary confusion), but we have already done so for our GenBank designations for trout GHRs. So, what were previously referred to as rainbow trout GHR1 and GHR2 (which were both in the type 2 GHR clade) (Very et al., 2005), are now rainbow trout GHR2a and rainbow trout GHR2b, respectively (cf. GenBank accession nos. NM_001124535 and NM_001124731). A similar scheme is proposed for the GHR1s. Whereas salmonids appear to have lost a gene following their 4R event and possess a single GHR1 (Fukada et al., 2005; Wallock et al., 2011) (GHR 1 is proposed to be used in preference to SLR so as to avoid confusion and to better represent the evolutionary origins of this gene), other species (e.g., Jian carp, Cyprinus carpio var. Jian) retained both GHR1 paralogs; these would be designated Jian carp GHR1a and Jian carp GHR1b.

4.3. Structure-function implications

Even with the clarified understanding of the evolutionary history of the GHR family, the possibility of novel functions through alternative transcripts of a gene or through heterodimerization be-

---

**Fig. 9.** Proposed phylogeny of the growth hormone receptor (GHR) family in vertebrates. The divergence of GHRs and prolactin receptors (PRLR) results from a series of gene duplication events over the course of vertebrate evolution. Subsequent duplication events in teleosts results in multiple types and subtypes of GHRs in this lineage.
tween different homologues in this family of receptors (or heterodimerization between different transcripts of the same or different homologue(s), for that matter) makes necessary further research regarding the molecular basis of functionality and how that functionality is transduced to diverse and integrated physiological effects of GH family peptides in animals.

Acknowledgments

We are grateful to Heather Bergan, Andrea Hanson, Naveen Lall, Lincoln Martin, Lindsey Norbeck, and Siddharth Satoshichandran for their technical assistance. We also thank Dr. Wayne Decatur for thoughtful discussions about modeling and for reviewing the manuscript. This work was supported by grants from the National Science Foundation, USA (IOS09200116 to M.A.S. and IOS0849569 to S.A.S.).

References


