Structure and expression of the mouse growth hormone receptor/growth hormone binding protein gene

J G Moffat, A Edens and F Talamantes
Department of Biology, University of California, Santa Cruz, California 95060, USA
(Requests for offprints should be addressed to F Talamantes)

ABSTRACT

The mouse growth hormone receptor/growth hormone-binding protein (GHR/BP) gene produces several distinct mRNA forms through alternative splicing, including mRNAs encoding the membrane-bound growth hormone receptor (GHR) and the soluble growth hormone-binding protein (GHBP). Transcripts are also heterogeneous in their 5' regions due to alternative selection of two major 5' untranslated region (5'UTR) sequences, designated L1 and L2. Here we report the cloning of all mouse GHR/BP coding exons as well as the exon encoding 5'UTR L2, the most widely expressed 5'UTR. The mouse GHR/GHBP gene contains 11 coding exons, 9 of which are homologous in size and sequence to human GHR exons 2–10. The two mouse exons that do not have homologs in the human gene are designated exons 4B and 8A. Exon 4B, located between exons 4 and 5, encodes an 8-amino acid segment of the ligand binding domain that is unique to mouse GHR and GHBP. Analysis by reverse transcriptase-polymerase chain reaction indicated that exon 4B is constitutively present in mouse GHR and GHBP mRNA. Exon 8A encodes the GHBP hydrophilic tail and 3'UTR sequence. 5'UTR L2 is encoded by a single exon located at least 27 kb upstream of exon 2 and at least 12 kb upstream of the exon encoding 5'UTR L1. The transcription start sites of UTR L2 were mapped and the 5' flanking region sequenced. The exon and proximal promoter region are GC rich, and share a high level of conservation with the equivalent exons in the sheep, bovine and human GHR genes. A CCAAT motif and several putative Sp1 motifs are present, and there is no TATA box. Homology between the mouse sequence and other species is limited to a region of 450 bp upstream of the exon due to the insertion of a fragment of a LINE-1 element upstream of the mouse L2 exon. Ribonuclease protection assays were used to confirm that 5'UTR L2 is widely expressed in multiple tissues and is the predominant form of transcript except in the liver during pregnancy, in which 5'UTR L1 is the major form.

Journal of Molecular Endocrinology (1999) 23, 33–44

INTRODUCTION

Growth hormone (GH) is a pituitary polypeptide hormone with multiple somatogenic and metabolic effects (reviewed by Harvey et al. 1995). Response to GH in target tissues is mediated by the growth hormone receptor (GHR), a single-pass transmembrane protein of the cytokine receptor superfamily (reviewed by Kelly et al. 1993). GHR receptor has been detected in practically all tissues examined (Harvey & Hull 1995), although the level of expression at the protein and mRNA levels varies widely, with the highest level of receptor being found in liver, fat and kidney. In serum, GH can be bound by growth hormone-binding protein (GHBP), a circulating glycoprotein that binds the hormone with affinity comparable to that of the receptor (reviewed by Barnard & Waters 1997).

In all species examined GHBP corresponds to the extracellular hormone-binding domain of GHR and is a product of the same gene, although the method of GHBP generation differs between species. In non-rodent mammals and chicken, generation of GHBP involves proteolytic cleavage of the full-length GHR protein. In rodents, however, distinct GHR- and GHBP-encoding mRNAs are generated by alternative splicing of a common primary transcript (reviewed in Edens & Talamantes 1998). The macaque is the only species identified to date in which GHBP can be produced by both alternative splicing and proteolytic cleavage (Martini et al. 1997), suggesting that further complexity remains to be uncovered. The relevant portions of the mouse
and rat GHR/BP genes have been cloned (Edens et al. 1994, Zhou et al. 1994, 1996). In both rodent species an alternatively-spliced exon (exon 8A) encodes a GHBP-specific hydrophilic C-terminal domain and the 3' untranslated region (UTR) of the GHBP mRNA. Mutually exclusive splicing of exon 8A or exon 8 to exon 7 thus generates the GHBP mRNA. Mutually exclusive splicing of exon containing two different 5' UTRs have been cloned in the mouse (Southard et al. 1995) and we have recently detected four other UTR sequences using 5' RACE (Moffat et al. 1998). Similarly two UTRs have been reported for GHR mRNA from chicken (Agarwal et al. 1994), and sheep (Adams 1995, O'Mahoney et al. 1994), three in the rabbit (Leung et al. 1987) and bovine (Hauser et al. 1990, Heap et al. 1996), five in the rat (Domene et al. 1995), and ten in the human (Pekhletsky et al. 1992, Schwartzbauer et al. 1998). Differences in hormonal, physiological, and tissue-specific expression of 5'UTRs suggest they are associated with different promoters (Adams 1995, Baumbach & Bingham 1995, Domene et al. 1995, Zhou et al. 1997, Lucy et al. 1998). We have previously reported that in the mouse two 5'UTRs, designated L1 and L2, show differential regulation in the liver and placenta during pregnancy (Southard et al. 1995). Expression of 5'UTR L1-containing GHR/BP mRNA is under the control of a liver-specific enhancer element (Menon et al. 1995), and pregnancy has been shown to induce a dramatic increase in transcription (Ilkbahar et al. 1999).

Splicing variants of the coding regions of GHR mRNA have also been reported. cDNA clones of human GHR have shown that exon 3, encoding 22 amino acid residues in the extracellular domain, can be included or excluded (Wickelgren et al. 1995, Zogopoulos et al. 1996). The sequence equivalent to exon 3 is constitutively present in mouse (Barnard et al. 1993) and absent in chicken (Agarwal et al. 1994) GHR/BP cDNA. The mouse cDNA sequence also contains an insertion equivalent to 8 amino acid residues compared with other species, at a position equivalent to the human exon 4/5 boundary (Godowski et al. 1989). Alternative splice acceptor site selection has also been detected in the exons encoding the intracellular domain of GHR, resulting in mRNAs potentially encoding C-terminally truncated proteins (Dastot et al. 1996).

The coding region of the GHR/BP gene has been cloned in the human (Godowski et al. 1989) and the chicken (Agarwal et al. 1994). The overall organization of these genes is typical of members of the cytokine receptor family. The human gene contains nine coding exons (designated exons 2–10, the numbering takes into account the presence of a 5'UTR) spanning 90 kb. Exons 2 through 7 encode the extracellular hormone-binding domain, exon 8 encodes the transmembrane domain, and exons 9 and 10 together encode the intracellular domain. The structure of the chicken GHR/BP gene is similar to that of the human gene (Agarwal et al. 1994).

To gain a more complete understanding of the regulation of GHR/BP gene expression and the generation of alternative transcripts we have cloned the complete mouse GHR/BP gene, including exons corresponding to all coding regions. The exon encoding 5'UTR L2 and its 5' flanking regions are also described. In addition, the expression pattern of the two mouse 5'UTRs is further characterized, showing that 5'UTR L2 is associated with a general promoter with little tissue specificity, while 5'UTR L1 is absent except for the previously reported pregnancy-specific hepatic upregulation.

**MATERIALS AND METHODS**

**Animals**

Swiss Webster mice were purchased from Simonsen Laboratories (Gilroy, CA, USA). All procedures for the care and use of animals were approved by the Chancellor's Animal Research Committee of the University of California, Santa Cruz.

**Construction of genomic library**

Genomic DNA was isolated from liver of Swiss Webster mice as described in Sambrook et al. (1989). DNA was further purified on a CsCl gradient and 30 µg were digested with BgIII (Promega, Madison, WI, USA) and dephosphorylated by incubating for 30 min at 37 °C with 0.5 units calf intestinal phosphatase. The genomic DNA was ligated into BamHI-cut EMBL3 vector DNA arms and packaged using the Packagene system (Promega) according to the manufacturer's instructions.

In addition, a mouse genomic library was purchased (Clontech, Palo Alto, CA, USA). This second library contained genomic DNA from mouse strain DBA/2J cloned into the BamHI site of the bacteriophage vector EMBL3 SP6/T7.

*Journal of Molecular Endocrinology* (1999) 23, 33–44
Library screening

Approximately 1 million recombinant bacteriophages were screened by transfer to Nytran membranes (Schleicher & Schuell, Keene, NH, USA) and hybridization to radiolabeled mouse GHR cDNA probes. Probes were generated by random priming of cDNA fragments in the presence of [α-32P]dATP (Sambrook et al. 1989). Hybridization of probes to the membranes was carried out overnight at 60 °C in 50 mM piperazine-N,N'-bis (2-ethanesulfonic acid) (pH 6.5), 50 mM sodium phosphate (pH 7.0), 100 mM NaCl, 1 mM EDTA, and 5% SDS.

Phage DNA analysis

Plaques that hybridized to the GHR probe were expanded and phage DNA was isolated.

Each clone was restriction-mapped using XhoI, XbaI, and EcoRI. Southern analysis was performed on all digestions to identify exon-containing restriction fragments, which were subcloned into pBluescript SK(+) (Stratagene, La Jolla, CA, USA) for sequencing. Sequencing was carried out using either Sequenase 2·0 (Amersham Pharmacia, Arlington Heights, IL, USA), or ABI Prism dye terminator reagents (Perkin Elmer, Foster City, CA, USA) with analysis on an ABI 310 sequence analyzer.

Reverse transcriptase-polymerase chain reaction (RT-PCR)

Mouse (Swiss Webster) and rat (Sprague Dawley) liver RNA (1 µg) were reverse-transcribed using oligo(dT) and Superscript II reverse transcriptase (Life Technologies, Gaithersberg, MD, USA) according to the manufacturer’s instructions. Forward and reverse primers complementary to sites in exon 4 (p4F, 5'-GACCCAGGATCTATTCA GC-3') and exon 5 (p5R, 5'-CAGGTTGCACAT TTCGTCAC-3') respectively, which are conserved between rat and mouse cDNA sequences, were used to amplify the exon 4/5 boundary region. PCR was carried out using 1 µl of a 1/100 dilution of reverse transcription product as template in a 20 µl reaction in 1 × PCR buffer (Promega), containing 0-5 units Taq polymerase (Promega), 2·5 mM MgCl2 and 0·2 µM of each primer. Amplification was carried out for 30 cycles and products were analyzed by electrophoresis on a 3% Metaphor (FMC, Rockland, ME, USA) gel.

S1 nuclease mapping

Total RNA was extracted from non-pregnant female Swiss Webster mouse liver using the method of Chomczynski and Sacchi (1987). Primer L2R1 (5'-AGTGCCGCCGCTCCCTAGAC-3') complementary to the 5' end of the UTR L2 sequence was end-labeled with [γ-32P]ATP, using T4 polynucleotide kinase (Promega) according to the manufacturer’s instructions. Single-stranded probe was generated by extending the primer with Taq polymerase for twenty cycles of PCR, using the 270 bp BamHI fragment encompassing the 5' flank of exon L2 as template, followed by gel purification. Probe (5 × 10⁴ c.p.m.) was ethanol precipitated with total mouse liver or yeast RNA (50 µg) and resuspended in 20 µl RPA hybridization buffer (Ambion, Austin, TX, USA), then incubated for 5 min at 75 °C and annealed at 42 °C for 16 h. Free probe was degraded by addition of 200 units S1 nuclease (Promega) in 200 µl S1 buffer for 30 min at 37 °C, followed by ethanol precipitation. Products were analyzed by electrophoresis on an 8% sequencing gel alongside a sequencing ladder generated using Sequenase (Amersham Pharmacia) with the same primer.

Ribonuclease protection assay

Total RNA was isolated from multiple tissues of 6-week-old non-pregnant and day 14 pregnant female and 8-week-old male Swiss Webster mice using a modification of the guanidine-phenol-chloroform method (Chomczynski & Sacchi 1987). Phenol-chloroform extractions were carried out twice and the RNA was subsequently re-precipitated twice with LiCl. Templates for probes containing mouse (m) GHR exons 2+3 and the 3' end of 5'UTRs L1 or L2 were generated by RT-PCR from mouse liver RNA and cloned into plasmid pGEM-T-easy (Promega). Antisense RNA probes were generated using T7 polymerase (Ambion) and gel-purified. Total RNA (10 or 50 µg as indicated) was annealed with 5 × 10⁴ c.p.m. of probe and ribonuclease protection assays (RPA) were carried out using the Hybspeed RPA kit (Ambion) according to the manufacturer’s instructions. Products were analyzed by electrophoresis on a 6% sequencing gel alongside size markers generated using Century marker templates (Ambion).

RESULTS

Isolation and characterization of clones

Six bacteriophage clones that hybridized to the mouse GHR cDNA probe were isolated from the Swiss Webster genomic library. A further clone containing the exon encoding 5'UTR L2 (exon L2) was isolated by screening the DBA/2J library with a
cDNA probe containing 5' UTR L2. Together, the clones represented 90 kb of genomic DNA and included all coding exons of the gene as well as exon L2. An eighth genomic lambda clone, containing the exon encoding 5' UTR L1, was previously cloned and characterized (Menon et al. 1995). Introns L2/L1, 2/3, 3/4, and 5/6 were not fully cloned and are therefore of undetermined lengths. Figure 1 shows the exon structure and restriction map of the gene. All splice sites conform to the canonical donor and acceptor consensus sequences. The exon/intron boundary sequences and sizes are summarized in Table 1.

The mouse GHR/BP gene contains eleven coding exons. Exons homologous to the nine human GHR/BP coding exons are present, and show conservation of splice boundaries and codon phasing patterns, whereas exons 4B and 8A do not have homologs in the human GHR/BP gene, as will be discussed below. Exon 2 encodes the final 9 base pairs of the 5'UTR sequence and the first 23 amino acids of the 24-amino acid signal sequence. Exons 3, 4, 4B, 5, 6, and 7 together encode 246 amino acids of the 249-amino acid hormone binding domain. Exon 8 encodes the final 3 amino acids of the extracellular domain, the 24-amino acid hydrophobic transmembrane domain, and the first 4 amino acids of the intracellular domain, while the alternative exon 8A encodes the GHBP hydrophilic tail and the 3'UTR of GHBP transcripts as has been described previously (Edens et al. 1994, Zhou et al. 1994). Exons 9 and 10 encode the intracellular domain of the receptor.

During cDNA cloning of the GHR mRNA the full-length 3'UTR region was not obtained; the cDNA was approximately 2 kb in size whereas Northern blotting showed that the GHR mRNA is 3·9-4·2 kb (Smith et al. 1988). Therefore a region of 2·5 kb 3' of the stop codon in exon 10 was sequenced, and a putative polyadenylation site was identified 2·0 kb downstream. This site is consistent with the predicted size of the GHR mRNA. The 3'UTR sequence contains several purine-rich (AUUUA) regions similar to elements known to regulate mRNA turnover (Hentze 1991). The functional significance of these sequences in the GHR mRNA has not been examined.

---

**Table 1. Exon-intron organization of mGHR/BP gene**

<table>
<thead>
<tr>
<th>Exon</th>
<th>Region encoded</th>
<th>Exon size (bp)</th>
<th>5' Boundary</th>
<th>3' Boundary</th>
<th>Intron size (kbp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>L2</td>
<td>5'UTR L2</td>
<td>232</td>
<td>gatgggTGGGAGAG..</td>
<td>..ACTCCGAG gtactg</td>
<td>&gt;12</td>
</tr>
<tr>
<td>L1</td>
<td>5'UTR L1</td>
<td>126</td>
<td>agatttATATACCC..</td>
<td>..TTGCAAAG gttaaa</td>
<td>15</td>
</tr>
<tr>
<td>2</td>
<td>Signal peptide</td>
<td>79</td>
<td>ttgcaag GTCTCGAG..</td>
<td>..AGT GAG G gtgaag</td>
<td>&gt;18</td>
</tr>
<tr>
<td>3</td>
<td>Extracellular</td>
<td>66</td>
<td>tttcaag CT ACA CC..</td>
<td>..GGG ACA A gtgaag</td>
<td>&gt;6</td>
</tr>
<tr>
<td>4</td>
<td>Extracellular</td>
<td>130</td>
<td>ctgcaag GT TCT TC..</td>
<td>..CT AAA AG gtgaag</td>
<td>1-5</td>
</tr>
<tr>
<td>4B</td>
<td>Extracellular</td>
<td>24</td>
<td>cttcaag G GAA AGC..</td>
<td>..CT GCA AG gtagt</td>
<td>6</td>
</tr>
<tr>
<td>5</td>
<td>Extracellular</td>
<td>173</td>
<td>tttcaag A ATT GCT..</td>
<td>..GAA ATA G gtgaagc</td>
<td>&gt;14</td>
</tr>
<tr>
<td>6</td>
<td>Extracellular</td>
<td>179</td>
<td>ttcagag TG CAA CC..</td>
<td>..G AA GTG gtgaag</td>
<td>5</td>
</tr>
<tr>
<td>7</td>
<td>Extracellular</td>
<td>169</td>
<td>tgagag ATG GGC C..</td>
<td>..GAA GAA G gtgaag</td>
<td>0-28</td>
</tr>
<tr>
<td>8A</td>
<td>GHBP tail and 3'UTR</td>
<td>161</td>
<td>cttcaag GA ACC AA..</td>
<td>..TTTTAAAAattaaaca</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>Transmembrane</td>
<td>91</td>
<td>tttaag AT ATC CA..</td>
<td>..AGCAA AG gtgaagq</td>
<td>4-5</td>
</tr>
<tr>
<td>9</td>
<td>GHR intracellular</td>
<td>70</td>
<td>tccaag G ATT AAG..</td>
<td>..T CTC AAG gtgaact</td>
<td>0-27</td>
</tr>
<tr>
<td>10</td>
<td>GHR intracellular and 3'UTR</td>
<td>3 (kbp)</td>
<td>ttctag GGA GGG A..</td>
<td>..AAGTTCTaataaa</td>
<td></td>
</tr>
</tbody>
</table>

Exon sequences are in capital letters, intron and flanking sequences are in lowercase letters. Splice donor and acceptor consensus residues and polyadenylation sequences are underlined.
A novel exon encodes an insert in the extracellular domain

Comparison of the mouse GHR/BP nucleotide and predicted amino acid sequences with those of other species reveals a novel 8-amino acid insertion (ESQRQAAR) in the extracellular domain. The site of this insertion corresponds to the junction between exons 4 and 5 in the human gene. The corresponding exons 4 and 5 in the mouse gene show equivalent boundaries and thus do not encode the insertion. Further analysis and sequencing of the intron between exons 4 and 5 identified a novel 24-nucleotide exon, which we have designated exon 4B encoding an insertion in the extracellular domain of mGHR. (A) Alignment of GHR amino acid sequences from the indicated species. Regions encoded by exons 4 and 5 (mouse and human genes), and exon 4B (mouse only) are indicated. Dark gray shading indicates positions of amino acid identity; light shading indicates conservative substitutions. Dashes indicate gaps introduced into sequences to maximize alignment. (B) Design of RT-PCR to test the presence or absence of exon 4B in GHR/BP mRNA, showing the expected product sizes for the two potential variants. Primers complementary to conserved sequences in exon 4 (p4F) and exon 5 (p5R) are indicated by the arrows. (C) Products generated by RT-PCR using RNA from rat liver and mouse kidney, liver, mammary gland (Mam. gland) and skeletal muscle (Skel. muscle) as indicated. Samples were electrophoresed on a 3% Metaphor agarose gel. Arrows indicate sizes of molecular weight standards (MW).
4B. Exon 4B is located 1.5 kb downstream of exon 4 and 6 kb upstream of exon 5. Figure 2A shows the exon/intron boundaries of exons 4, 4B, and 5, and a cross-species comparison of the amino acid sequence in this region of the GH-binding domain.

Given that the reading frame phasing is conserved with or without exon 4B, we investigated whether there was evidence for alternative splicing leading to the exclusion of this region, similar to that reported for exon 3 in humans. GHR/BP mRNA from mouse liver, kidney, mammary gland, and muscle, and rat liver was analyzed by RT-PCR using primers flanking exon 4B (Fig. 2B). As shown in Fig. 2C, exon 4B appears to be constitutively present in mouse GHR/BP mRNA but absent in rat liver GHR/BP mRNA, consistent with previous cDNA cloning studies (Mathews et al. 1989, Smith et al. 1989).

Sequence and transcription start site of 5'UTR-L2 promoter region

Two 5'UTR sequences, L1 and L2, have previously been cloned in mouse GHR/BP cDNAs by this laboratory (Southard et al. 1995). Analysis of a genomic clone containing 5'UTR L1-encoding exon (exon L1) has been described (Menon et al. 1995). We have isolated and characterized a genomic clone (º66B) containing the 5'UTR L2 exon and its putative promoter. No overlapping clones linking the 5'UTR-containing clones to each other or to exon 2 were found; furthermore, the 5' end of the genomic clone containing exon L1 was found to be part of a LINE-1 retroposon (not shown), which would greatly hinder attempts to use genome walking to clone the upstream region. Long-distance genomic

---

**Figure 4.** Sequence of 5'UTR L2 and the putative proximal promoter region. Lower case letters indicate splice donor site and intron residues, exon sequence is in upper case. The major transcription start site as determined by S1 nuclease mapping extension is numbered as residue +1. Major and minor transcription start sites are indicated by asterisks. The polyadenine tract corresponding to the 3' end of the LINE-1 fragment is underlined and labeled. Boxed sequence indicates the conserved CCAAT. Putative Sp1 binding motifs and initiator (Inr) element are underlined and labeled. The conserved upstream open reading frame is underlined and labeled and the splice donor consensus sequence at the 3' end of the exon is shown in bold.

---
PCR was thus used to determine the order and organization of these exons. PCR amplification using a forward primer complementary to exon L1 and a reverse primer complementary to exon 2 produced a 15 kb product. The identity of the 15 kb genomic PCR product was confirmed by restriction mapping and comparison with the L1 and exon 2-containing lambda clones (data not shown). Comparison of restriction maps of the genomic PCR product with λ66B confirmed that exon L2 does not lie within this region. Repeated attempts to amplify a genomic PCR product using primers complementary to sequences downstream of L2 and upstream of L1 were not successful, suggesting that the distance between exons L1 and L2 may be significantly greater than the minimum of 12 kb present in the non-overlapping lambda clones. This organization of alternative promoters and 5’UTR-encoding exons is similar to that reported for the sheep (O’Mahoney et al. 1994, Adams 1995), in which the liver-specific promoter and exon 1A (partially homologous to mouse L1) is located 17 kb upstream of exon 2 and the homolog of the mouse L2 promoter and exon, designated 1B, is a large (>8 kb) but unknown distance further upstream.

The transcription start site (tss) of exon L2 were determined by S1 nuclease mapping. A primer complementary to a sequence near the 5’ end of the longest L2 cDNA was used to generate a single-stranded end-labeled DNA probe which was hybridized to total mouse liver RNA prior to treatment with S1 nuclease. Products were electrophoresed alongside a sequencing ladder generated using the same primer with cloned genomic DNA, as shown in Fig. 3. The protected fragments, corresponding to the 5’ end mRNAs, identify multiple transcription start sites at positions between 200 and 230 bases upstream of the 3’ end of the exon. The major transcription start site, designated as position +1 in the sequence, was 230 bases upstream of the splice donor site. A minor product was also detectable, upon longer exposure, at a position 30 bp upstream of this position, which corresponds to the longest cDNA product upstream of exon 2 and the homolog of the mouse L2 promoter and exon, designated 1B, is a large (>8 kb) but unknown distance further upstream.

The transcription start sites (tss) of exon L2 were determined by S1 nuclease mapping. A primer complementary to a sequence near the 5’ end of the longest L2 cDNA was used to generate a single-stranded end-labeled DNA probe which was hybridized to total mouse liver RNA prior to treatment with S1 nuclease. Products were electrophoresed alongside a sequencing ladder generated using the same primer with cloned genomic DNA, as shown in Fig. 3. The protected fragments, corresponding to the 5’ end mRNAs, identify multiple transcription start sites at positions between 200 and 230 bases upstream of the 3’ end of the exon. The major transcription start site, designated as position +1 in the sequence, was 230 bases upstream of the splice donor site. A minor product was also detectable, upon longer exposure, at a position 30 bp upstream of this position, which corresponds to the longest cDNA product upstream of exon 2 and the homolog of the mouse L2 promoter and exon, designated 1B, is a large (>8 kb) but unknown distance further upstream.
previously cloned (Southard et al. 1995). The sequence CCAGTTT, which matches the consensus for the transcriptional initiation motif Inr (YYANT/AYY), commonly found in TATA-less promoters (Javahery et al. 1994), is present immediately upstream of the most distal transcription start site.

Contiguous restriction fragments spanning from 4-0 kb upstream of exon L2 to 0-5 kb downstream were sequenced. The sequence of the 5′ UTR L2 exon and proximal 5′ region is shown in Fig. 4.

Computer analysis of this sequence for transcription factor binding consensus sequences indicated that four potential binding sites exist for the constitutive transcription factor Sp1. A CCAAT motif is also present, and is conserved between species (see Fig. 5). The putative promoter and UTR sequences show several other interesting features which are presently of unknown significance. There is a very high GC content (70%), as well as a pyrimidine-rich region composed predominantly of T(C3–4) repeats. There is also, as previously reported (Southard

\[ \text{A.} \]

<table>
<thead>
<tr>
<th>L1 Probe:</th>
<th>5′ UTR-L1</th>
<th>exon 2+3</th>
<th>T7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protected fragments:</td>
<td>+ L1 235 nt</td>
<td>- L1 138 nt</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>L2 Probe:</th>
<th>5′ UTR-L2</th>
<th>exon 2+3</th>
<th>T7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protected fragments:</td>
<td>+ L2 169 nt</td>
<td>- L2 138 nt</td>
<td></td>
</tr>
</tbody>
</table>

\[ \text{B.} \]

\[ \text{C.} \]

**Figure 6.** Expression of 5′ UTRs L1 and L2 in adult mouse tissues. (A) Templates used to generate antisense RNA probes using T7 polymerase. The sizes of the expected fragments protected by each mRNA variant are indicated. (B) RPAs carried out using the 5′ UTR L1 probe. LP and cyclophilin probes were hybridized with 10 µg total RNA from the indicated mouse tissues (or 50 µg in the case of testis, ovary and brain RNA) and 50 µg yeast RNA (lane 3). Preg. Liver and NP Liver indicate liver RNA from day 14 pregnant and non-pregnant mice respectively. Undigested probes are shown in lane 2 next to the size markers. Bands corresponding to the 235-base L1-containing fragment, the 138-base non-L1 fragment and the 103-base cyclophilin fragment (C) are arrowed. (C) RPAs carried out as for (B) using the L2 probe. Arrows indicate the 169-base L2-containing and 138-base non-L2 fragments and the cyclophilin loading control (C).
et al. 1995), a short open reading frame present in the UTR sequence. As shown in Fig. 5, the mouse sequence exhibits a high degree of homology with the corresponding regions from the human, bovine and sheep sequences (due to the very high level of homology between sheep and bovine sequences the bovine has been omitted from the alignment). However, this homology ceases at a polyadenine sequence position −480 bases from the tss, which corresponds to the 3’ end of a LINE-1 retroposon element. Homology to LINE-1 sequence extends upstream for 320 bases. The inter-species homology does not resume upstream of this insertion.

General expression of 5’UTR L2 and restricted expression of 5’UTR L1

The differential usage of 5’UTR L1 and L2 promoters has previously only been characterized in liver from non-pregnant mice and liver and placenta from pregnant mice (Southard et al. 1995). 5’UTR L2 was found to be the predominant form of both GHR and GHBP mRNA in placenta and non-pregnant liver, whereas UTR L1 was predominant in the pregnant liver. In other species the UTR L2 homolog has been reported to be the predominant form in extrahepatic tissues (Adams 1995, Domene et al. 1995, Lucy et al. 1998). To further test the hypothesis that 5’UTR L2 is associated with a non-tissue specific constitutive promoter, the presence of each of the 5’UTR sequences in GHR/BP mRNA from multiple tissues was examined.

Ribonuclease protection assays were carried out using two probes illustrated in Fig. 6A. Each probe contains a common region of GHR/BP cDNA (the probes used do not distinguish between GHBP and GHR mRNA) corresponding to exons 2 and 3, which results in a protected fragment of 138 bases, along with the 3’ portion of either 5’UTR L1 or L2. For a given probe, the relative amounts of long and short protected fragments indicates the proportion of mRNA in each tissue which contains the particular 5’UTR. GHR/BP mRNA was detectable in all the tissues examined and the relative levels are in agreement with previous reports of GHR/BP protein and mRNA levels in mouse and other species. The data for the expression in liver of pregnant and non-pregnant mice is consistent with previous observations, other tissues including adipose and muscle have been shown to have no gestational upregulation of GHR/BP (Ilkbahar et al. 1995). The highest overall level of expression occurred in liver from pregnant mice. Figure 6B shows that in hepatic RNA from pregnant mice, mRNA containing 5’UTR L1 accounts for the elevated level of mRNA compared with hepatic RNA from non-pregnant animals, whereas the 138 nt band is unchanged. These data are consistent with our observations of a liver-specific upregulation of transcription from the L1 promoter during pregnancy (Southard et al. 1995, Ilkbahar et al. 1999). In all other tissues L1-containing mRNA is undetectable, except for a faint band detectable in the non-pregnant liver sample upon longer exposure.

The use of the 5’UTR L2 containing probe (Fig. 6C) demonstrates that liver from pregnant mice is the only tissue in which a large amount of non-L2-containing mRNA, detectable as a strong band of 138 nt, is present. We can infer that this corresponds to the L1-containing mRNA detected in Fig. 5B. In all other tissues examined, L2-containing mRNA is present in great excess of non-L2 species. Quantitation of relative intensities of the 169 nt and 138 nt bands by phosphorimager indicated that at least 95% of GHR/BP mRNA in these tissues contains 5’UTR L2 (data not shown). It is noteworthy, however, that a band corresponding to non-L2-containing mRNA is present in these tissues, despite the absence of detectable L1-containing mRNA. This is consistent with our recent observation of the occurrence of other 5’UTR sequences in mouse GHR/BP mRNA (Moffat et al. 1998).

DISCUSSION

The mouse GHR/BP gene is composed of 11 coding exons and at least two alternative first exons, which encode 5’UTR sequences. Nine of the coding exons are homologous in size and sequence to the 9 coding exons of the human GHR gene. The two coding exons that do not have homologs in the human GHR/BP gene have been designated exons 4B and 8A. Overall this structure is typical of members of the class I cytokine receptor family.

Aside from the previously characterized alternative exon 8A, encoding the 3’ terminus of the GHBP mRNA, the other difference from the human gene is the presence of exon 4B, encoding an 8-amino acid segment unique to the mouse GHR/BP hormone-binding domain. Although no investigations have directly addressed the functional significance of the 8-amino acid segment, the X-ray crystal structure of the human GHBP/GH complex shows that the site of the insertion is removed from the residues which interact with the GH molecule (De Vos et al. 1992). Furthermore, this region corresponds to a loop extending into the solvent from between strands of the beta sheet forming the hormone binding site, and the residues forming
this loop are absent from the crystal structure, signifying that this loop is disordered. Studies of the human GHR, however, have shown that alanine mutagenesis of amino acids at these sites reduces GH-binding by 2- to 3-fold (Bass et al. 1991). Together, these results suggest that the region encoded by the exon 4/exon 5 splice junction, while not directly involved in binding GH, may be important for proper folding of the GHR extracellular domain. Comparison of this region of sequence with GHR from ruminants and the chicken show that there is considerable variation in the number of amino acids in this loop region, implying that it is a structurally and evolutionarily plastic region.

The exon encoding 5'UTR L2 and its 5' flanking region were characterized in the present study. Sequence comparison reveals that exon L2 and the upstream region are highly homologous to exons V2 and 1B and their 5' proximal regions in the human (Rubtsov 1997) and sheep and bovine GHR genes (Adams 1995, Heap et al. 1996) respectively. The promoters in all three species lack a TATA box but have multiple GC-rich elements which are potential binding sites for the transcription factor Sp1. Multiple transcription start sites, such as detected in this work, are typical of promoters lacking a TATA box. The transcription start sites, as mapped by S1 nuclease protection, are located within 50 bases upstream of the ATG codon of the putative open reading frame, and are near a potential Inr motif. The reported transcription start sites in the sheep promoter are located further upstream, near the potential CCAAT motif, than the sites reported here for the mouse gene, while the transcription start site(s) for human exon V2 remain to be determined. Despite the discrepancy between the mouse and sheep transcription start sites in an otherwise very highly conserved sequence, the start sites reported here are consistent with the size of the mouse GHBP mRNA (1200–1300 nt) detected by Northern blotting (Smith et al. 1988).

The high degree of homology between the mouse L2 promoter and the sheep, bovine and human sequences is abolished 480 bases upstream of the mouse transcription start site by the insertion of a fragment of a LINE-1 mobile DNA element. Full length LINE-1 sequences are 6–7 kb in length and contain a promoter and two open reading frames (Fanning & Singer 1987). The partial LINE-1 insert found upstream of exon L2 contains only the 3' 320 bp of the full LINE-1 sequence, and therefore would not be expected to contain any of the promoter elements found in the 5' regions of LINE-1 sequences. Homology with the sheep and human promoters does not resume upstream of the insertion, although the sheep, human and bovine sequences continue to have a high degree of homology to each other for at least another 1 kb. Therefore, the divergence of the mouse sequence from the other species is not solely due to insertion of the LINE-1 fragment. The genomic clone containing exon L2 and the upstream LINE-1 fragment has only been cloned from a genomic library prepared from DBA/2J strain mice, and we have not yet examined the corresponding region from other strains of mice to determine whether the insertion is strain-specific. This possibility seems unlikely, however, as the few reported LINE-1 insertion events which are sufficiently recent to be strain-specific (e.g. Boccaccio et al. 1990) are characterized by an uninterrupted 3' polyadenine sequence corresponding to the poly-A tail of the transcript, and by short direct repeats flanking the insertion site. These regions have apparently undergone multiple mutations in the L2 upstream LINE-1, suggesting that this insertion is unlikely to have occurred within the timeframe of establishment of inbred strains of mice.

Previous studies showed that 5'UTR L1 is upregulated in the liver during pregnancy, while 5'UTR L2 predominates in the placenta and in the liver of non-pregnant mice (Southard et al. 1995). However detection of the different 5'UTR-containing transcripts was carried out by Northern blotting, which did not allow estimation of the relative contributions of each of the 5'UTRs to the total pool of GHR/BP mRNA, nor were other tissues examined. The ribonuclease protection assays presented in the current study allow a direct estimation of the relative levels of 5'UTR L1- and L2-containing mRNA in the previously studied tissues as well as in other GHR/BP-expressing tissues. Demonstration that 5'UTR L2 is the predominant form of the message in all tissues examined other than pregnant liver is consistent with features typical of many housekeeping promoters, including the absence of a TATA box and the presence of GC-rich Sp1 like elements, as well as a high GC content (Javahery et al. 1994). This pattern of expression is similar to that previously observed for the liver-specific (1A) and general (1B) 5'UTRs of sheep and bovine GHR (Adams 1995, Lucy et al. 1998), with the exception that in the mouse the liver-specific L1 promoter is even more restricted in its activity, with significant levels of L1-containing transcripts being present only in the liver during pregnancy. In spite of the wide distribution of L2-containing GHR/BP expression and the similarity of the L2 promoter to typical
constitutive promoters it must be noted that the level of expression varies widely between different tissues, implying that there are tissue-specific as well as housekeeping regulatory elements in the L2 promoter. Furthermore, the conserved GC-rich nature, the likelihood of significant stable secondary structure, and the conserved upstream open reading frame of this 5′UTR suggest that it may play a role in post-transcriptional regulation of GHR/BP production. The studies reported here will facilitate further examination of these questions.

ACKNOWLEDGEMENTS

We gratefully acknowledge Dr Linda Ogren for her advice and help in preparing this manuscript, Huy Dao for technical assistance, and Yonca Ilkbahar and Bernardo Contreras for discussions. We thank Dr Ram Menon, Childrens Hospital of Pittsburgh, Pennsylvania, for providing the genomic clone containing L1, and for helpful discussions. This work was supported by NIH grants CA71590, HD14966 and GM08132 to FT.

REFERENCES


Baumbach WR & Bingham B 1995 One class of growth hormone (GH) receptor and binding protein messenger ribonucleic acid in rat liver, GHR1, is sexually dimorphic and regulated by GH. Endocrinology 136 749–760.


Domenech HM, Castorla F, Werner HI, Roberts CP & Leroith D 1995 Rat growth hormone receptor/growth hormone-binding protein mRNAs with divergent 5′-untranslated regions are expressed in a tissue-specific manner. DNA and Cell Biology 14 195–204.


Edens A, Southard JN & Talamantes F 1994 Mouse growth hormone-binding protein and growth hormone receptor transcripts are produced from a single gene by alternative splicing. Endocrinology 135 2802–2805.


Hauser SD, McGrath MF, Collier RJ & Krivi GG 1990 Cloning and in vivo expression of bovine growth hormone receptor mRNA. Molecular and Cellular Endocrinology 72 187–200.


Pekhletsky RI, Chernov BK & Rubtsov PM 1992 Variants of the 5'-untranslated sequence of human growth hormone receptor mRNA. *Molecular and Cellular Endocrinology* **90** 103–109.

Rubtsov PM 1997 *H. sapiens* growth hormone receptor gene, 5'-flanking region. Direct submission Genbank accession no. AJ002175.


Zhou Y, He L & Kopchick JJ 1994 An exon encoding the mouse growth hormone binding protein (mGHBP) carboxy terminus is located between exon 7 and 8 of the mouse growth hormone receptor gene. *Receptor* **4** 223–227.


RECEIVED 4 November 1998

NOTE ADDED IN PROOF

Sequences have been deposited in the Genbank database, accession numbers AF120480-AF120480.