Structure of the glucocorticoid receptor (NR3C1) gene 5’ untranslated region: identification, and tissue distribution of multiple new human exon 1

Jonathan D Turner1 and Claude P Muller1,2
1Institute of Immunology, Laboratoire National de Santé, 20A rue Auguste Lumière, L-1011, Luxembourg
2Department of Immunology, Graduate School of Psychobiology, University of Trier, D-54290, Germany

Abstract

The 5’ untranslated region (UTR) of the glucocorticoid receptor (GR) plays a key role in determining tissue-specific expression and protein isoforms. Analysis of the 5’ UTR of the human GR (hGR) has revealed 11 splice variants of the hGR exon 1, based on seven exon 1s, four of which (1-D to 1-F and 1-H) were previously unknown. All of the exon 1 variants have unique splice donor sites and share a common exon 2 splice acceptor site. Due to an upstream in-frame TGA stop codon the predicted translation from all splice variants is identical. The four new exon 1s show remarkable similarity with their rat homologues. Exon 1-D starts and finishes 17 and 36 bp upstream of the corresponding ends of the rat exon 14. Exon 1-E is only 6 bp longer than its homologue exon 15. Exon 1-F contains two short inserts of 11 and 6 bp when compared with the rat 17. 1-H is 18 bp longer than the corresponding rat 111. In addition to these new exons, we found that the human exon 1-C occurs as three distinct splice variants, covering the region homologous to the rat exons 1a and 1p. All of the alternative hGR exons 1s presented here were found to be transcribed in human tissue. The human hippocampus expresses mRNA of all the exon 1 variants, while the expression of the other exon 1s seems to be tissue specific. While exon 1-D is only in the hippocampus, exons 1-E and 1-F are also detected in the immune system, and exon 1-H additionally in the liver, lung and smooth muscle. The 5’ region of the hGR is more complex than previously thought, and we suggest that each of these untranslated first exons have a distinct proximal promoter region, providing additional depth to the mechanisms available for tissue-specific expression of the hGR isoforms.

Journal of Molecular Endocrinology (2005) 35, 283–292

Introduction

Glucocorticoids play an important role in regulating the intermediary metabolism, in controlling cell differentiation, nervous and immune system functions, and in maintaining homeostasis after stress. These diverse activities are mediated by the type II glucocorticoid receptor (GR, OMIM +138040; official symbol is NR3C1 for nuclear receptor subfamily 3, group C, member 1), a key member of the nuclear receptor family. For these important functions the fine-tuning of GR levels is critical both in vitro (Vanderbilt et al. 1987) and in vivo. It is known that GR is expressed in virtually all cell types, but levels of both GR mRNA and protein vary considerably between tissues. GR expression is also subjected to a significant level of plasticity in some tissues such as the hippocampus (Herman et al. 1989). In transgenic mice a 30–50% reduced tissue level of GR showed major neuroendocrine, metabolic and immunological abnormalities (Pepin et al. 1992, King et al. 1995). Thus rodent GR activity relies on a tight control of its expression.

In eukaryotic cells gene expression is controlled by a variety of mechanisms, both at transcriptional and translational levels. Chromatin condensation, initiation, DNA methylation, alternative RNA splicing and mRNA stability are important transcriptional factors in many genes and tissues. In rodents the regulation of GR levels occurs almost exclusively at the transcriptional level (Strahle et al. 1992). This observation that has been recently confirmed also for the human GR (hGR) (Pedersen et al. 2004), but, so far the hGR gene seems much less complex than the rat GR gene. The hGR gene covers a region of more than 80 kb within chromosome 5, containing eight coding exons (exons 2–9) and alternative 5’-non-coding exon 1s (Zong et al. 1990, Encio & Detera-Wadleigh 1991, Breslin & Vedeckis 1998, Breslin et al. 2001). Only three alternative (heterogenous) non-coding exon 1s, and two of the corresponding promoter regions have been described for the hGR (Breslin & Vedeckis 1998, Breslin et al. 2001). Careful examination of the gene sequence and the known hGR transcription initiation sites have shown that each previously reported
initiation site has its own promoter region (Barrett et al. 1996, Wei & Vedeckis 1997, Breslin & Vedeckis 1998, Breslin et al. 2001, Nunez & Vedeckis 2002, Geng & Vedeckis 2004). Multiple independent promoter regions producing the same protein by splicing upstream of the translation initiation site, partially explain tissue-specific GR expression (Nunez & Vedeckis 2002, Pedersen & Vedeckis 2003, Geng & Vedeckis 2004). The presence of a glucocorticoid response element in only one of the two currently reported hGR promoter regions helps to explain why GR expression is up-regulated in some and down- or unregulated in other tissues in response to endogenous and exogenous glucocorticoids (Zong et al. 1990, Strahle et al. 1992).

The hGR is a member of a particular subset of the nuclear receptors, the steroid hormone receptors (oestrogen, androgen and progesterone receptors). Oestrogen receptor α (ERα) has a gene structure similar to the hGR, consisting of eight coding exons and heterogeneous non-coding first exons. For hERα, eight untranslated first exons have been described (E1, F, E2, T1, T2, D, C and B), and these are expressed in a tissue-specific manner. These transcription start sites each paired with its own promoter region seem to correlate with tissue-specific expression of the ERα. The presence of a similar transcriptional control mechanism for the hGR can be expected from the structure of the 5’ region of the rat GR mRNA where 11 different heterogeneous rat exon 1 variants have been identified (McCormick et al. 2000).

Here we investigated the 5’ UTR of the hGR and identified all of the corresponding human exon 1s including four novel exon 1s, three splice variants with their 5’ and 3’ boundaries, and their tissue distribution including cells of the immune system.

Materials and methods

Identification of novel exon 1s

Initial alignments of the genomic DNA of the human chromosome 5 (July 2004, University of California Santa Cruz (UCSC)) and of the rat (AJ271870, National Center for Biotechnology Information (NCBI)) upstream of the GR exon 2 was performed using Vector NTi (Invitrogen). The known GR exon 1s of Rattus norvegicus were localised by their published 3’ end sequences. Forward PCR primers (Table 1) were developed on the basis of the corresponding human exon 1s. These primers were paired with a common reverse primer located in the human exon 2. Amplification of cDNA by PCR was performed using 20 mM Tris–HCl (pH 8-4), 50 mM KCl, 200 mM dNTPs and 2·5 U Platinum TaqDNA Polymerase (Invitrogen) (35 cycles: 96 °C, 20 s; annealing, 20 s; 72 °C, 20 s), followed by 10 min at 72 °C. Annealing temperatures, primers and MgCl₂ concentrations are shown in Table 1. Oligo-dT-primed universal human cDNA (BD Clontech, Belgium) was used as a template, providing cDNAs from a variety of human tissues.

Identification of the 3’ boundary

The above PCR products were sequenced to identify the 3’ boundaries of novel exon 1s and to confirm those previously reported. For this purpose, fresh PCR products were cloned into the pCR4-TOPO vector (Invitrogen) and used to transform Top10 electrocompetent Escherichia coli (Invitrogen). Colonies were screened using M13 primers. M13 PCR products of the correct size were purified with spin columns (Life Technologies, Paisley, UK), and sequenced using 100 nM M13 primers (Table 1) and the BigDye 3·1 terminator cycle sequencing reagent (Applied Biosystems, Nieuwerkerk, The Netherlands). After 35 cycles (96 °C, 20 s; 50 °C, 20 s; 72 °C, 2 min) the reaction was completed by incubating for 10 min at 72 °C. The products were precipitated with ethanol and sodium acetate (0·3 M). Sequencing was performed on either an ABI PRISM 377 or ABI 3130 sequencer (Applied Biosystems).

Cell preparations and biopsy samples

Human peripheral blood mononuclear cells (PBMCs) were purified from a buffy coat bag (Red Cross, Luxembourg) using Leucosep tubes (Greiner Bio-One, Frickenhausen, Germany) and Ficol-Isopaque (Amersham) according to the manufacturers’ instructions. After three washes in Hanks balanced salt solution (HBSS) (Biowhittaker, Verviers, Belgium) cells were resuspended in PBS, 0·5% BSA and 2 mM EDTA. Cellular subsets were purified by positive selection with monoclonal antibodies coupled to magnetic beads (Miltenyi Biotech GmbH, Cologne, Germany). Briefly, cells were incubated for 15 min with magnetic beads carrying the appropriate monoclonal antibodies in PBS containing 2% BSA, 20 mM EDTA and subsequently separated in the midiMacs system using a positive separation column (LS). Cell subset purity and identity was confirmed by flow cytometry (EPICS ELITE ESP, Coulter, Namur, Belgium). A purity of 98–99·5% was observed (data not shown).

Anonymous human biopsies were obtained from surgically removed organ tissues from the Clinic Sainte Therese, Luxembourg, following national and institutional ethical guidelines. The following tissues were obtained: thyroid gland, salivary gland, liver, prostate, skin, sub-cutaneous adipose tissue, lung and heart.
muscle. Purified immune cells included CD4+ T cells, CD8+ T cells, CD14+ monocytes, CD19+ B cells and CD11c+/CD123+ (BDCA-4+) plasmacytoid dendritic cells. Hippocampal cDNA (dT16 primed) was obtained from Biochain (Gentaur, Brussels, Belgium).

Isolation of mRNA, and reverse transcription

Briefly, 10⁶ purified cells or 10 mg of tissue were lysed in lysis buffer (Miltényi); after removal of debris, samples were incubated with oligo-dT-labelled magnetic beads (Miltényi). mRNA was subsequently isolated using µMacs magnetic columns. First-strand synthesis of total cDNA was carried out at 50 °C for 60 min using 200 U SuperScript III RT and 2·5 μM dT16 primer in a 40 μl reaction containing 250 mM Tris–HCl (pH 8·3, at room temperature), 375 mM KCl, 15 mM MgCl₂, 10 mM dithiothreitol and 500 μM dNTPs.

5′-rapid amplification of cDNA ends (RACE) PCR

The first strand of specific GR cDNA was reverse transcribed for 1 h at 42 °C using 8 U/ml Moloney murine leukemia virus (MMLV) reverse transcriptase and 100 nM exon 2-specific reverse primer (2Det, Table 1) in a 25 μl reaction containing 20 mM Tris–HCl (pH 8·4), 50 mM KCl, 3 mM MgCl₂, 20 mM dithiothreitol and 400 μM dNTPs. After RNase H (Invitrogen) treatment for 10 min (37 °C), a poly-C tail was added to the 3′ end of the cDNA. To a 25 μl reaction containing 20 mM Tris–HCl (pH 8·4), 50 mM KCl, 1·5 mM MgCl₂, 200 mM dCTP and 0·4 U/ml terminal deoxynucleotidyl transferase, 1 μl cDNA was added. Poly-C-tailed cDNA (1 μl) was subsequently amplified using a nested PCR reaction. First-round touchdown PCR was performed with 400 nM abridged anchor primer (AAP, Table 1, Invitrogen), 400 nM exon 2-specific reverse primer (2 RACE RC, Table 1),

Table 1 PCR primers and their associated reaction conditions

<table>
<thead>
<tr>
<th>Detection forward primers</th>
<th>Sequence</th>
<th>Primer pair</th>
<th>[Mg²⁺] (mM)</th>
<th>Tₘ (°C)</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-BDet</td>
<td>GCCGGGACGCGA CCTC</td>
<td>2Det</td>
<td>2</td>
<td>55</td>
<td>0·5 mM primers</td>
</tr>
<tr>
<td>1g hom</td>
<td>TGGCTTCGCAACGTAAGTT</td>
<td>2Det</td>
<td>2</td>
<td>53</td>
<td>1g homologue</td>
</tr>
<tr>
<td>110 hom</td>
<td>CGAGCCTCGTTGTCTGTGAC</td>
<td>2Det</td>
<td>2</td>
<td>55</td>
<td>110 homologue</td>
</tr>
<tr>
<td>1-CtotDet</td>
<td>GCTCCTGCGAAGATTTGAT</td>
<td>2Det</td>
<td>2</td>
<td>55</td>
<td>Detects 1-C1, 1-C2 and 1-C3</td>
</tr>
<tr>
<td>1-C2Det</td>
<td>TCTGCTGTGAGGATTCTGC</td>
<td>2Det</td>
<td>2</td>
<td>55</td>
<td>Detects 1-C2 only</td>
</tr>
<tr>
<td>1-C3Det</td>
<td>CTATAATAGGGCTTCCC</td>
<td>2Det</td>
<td>2</td>
<td>53</td>
<td>0·5 mM primers</td>
</tr>
<tr>
<td>1-DDet</td>
<td>ACAACCTTTCAGAGTGC</td>
<td>2DET</td>
<td>3</td>
<td>56</td>
<td>Detects 1-C3 only</td>
</tr>
<tr>
<td>1-Dext</td>
<td>GGCTGTTGGCCAGGCCGC</td>
<td>2DET</td>
<td>2</td>
<td>58</td>
<td>Detection in tissue panel</td>
</tr>
<tr>
<td>1-FDet</td>
<td>GTGAGCGAGAAAAAAGACTGG</td>
<td>2DET</td>
<td>3</td>
<td>55</td>
<td>Primer extension assay</td>
</tr>
<tr>
<td>1-G</td>
<td>TGAGGGGCGAGGCACCTGGA</td>
<td>2DET</td>
<td>4</td>
<td>59</td>
<td>No product detected under conditions tested</td>
</tr>
</tbody>
</table>

GR(total) quantification

| 2Fwd                      | GTTGATTTTCCAAAGG | 2Rev         | 3           | 48      | |
| 2Rev                      | CAATGCTTTTCCTCA | 2Fwd         | 3           | 48      | |

Conserved primers

| 2Det                      | CAGTGCGATCCTGA AACTCTTG | N/A | For detection |
| AAP                       | GGCACGCCGCGTCACTAGTG | N/A | For 5′ RACE |

Sequencing primers

| M13Fwd                    | GTAAAACGACGGCCAG | M13Rev      | 2           | 55      | |
| M13Rev                    | CAGGAAACAGCTATGAC | M13Fwd      | 2           | 55      | |

| Tₘ, annealing temperature in PCR; TD, touchdown PCR using five cycles annealing at 68 °C, five cycles at 62 °C, and 25 cycles at 55 °C; Fwd, forward; Rev, reverse. |
20 mM Tris HCl (pH 8·4), 50 mM KCl, 1·5 mM MgCl2, 200 mM dNTPs and 2·5 U Platinum Taq DNA Polymerase (35 cycles: denaturation, 96°C, 20 s; annealing, 68°C cycles 1–5, 62°C cycles 10–35, all cycles for 20 s; elongation, 72°C, 20 s), with a final elongation for 10 min at 72°C. Nested exon 1-specific 5′-RACE products were subsequently generated using the protocol above and the reverse primers shown in Table 1. Fresh PCR products were cloned into the pCR4-TOPO vector (Invitrogen) and subsequently sequenced using M13 primers as described above. The nucleotide sequence data (Table 2) were deposited with the EMBL and GenBank Nucleotide Sequence Databases.

**Tissue distribution of exon 1 variants**

Tissue distribution of the exon 1 variants was determined by classical PCR. Reaction conditions and primer sequences (1-BDet, 1-C2Det, 1-C3Det, 1-CtotDet, 1-DDet, 1-FDet, 1-HDet and the common reverse primer 2Det) are listed in Table 1. To ensure that the hGR was detectable in each sample, a PCR of the coding region of exon 2 (2Fwd and 2Rev) was included in each sample.

**Results**

**Prediction and detection of human homologues to the rat exon 1s**

To investigate the 5′ heterogeneity of the hGR mRNA the alternative start sites of transcription were localised in the published 4·9 kbp promoter and first exon-rich region of the rat genomic DNA sequence. This was aligned with the relevant region of human chromosome 5. The alignment of regions known to encode the rat alternative exon 1s with their human homologues is shown in Fig. 1. To confirm the existence of these homologous exon 1s, PCR forward primers within each of these exons were combined with a common reverse primer from the 3′ end of exon 2. The location of the primers used is illustrated in Fig. 1. RT-PCR performed on universal human cDNA representing a broad range of expressed genes revealed the presence of eight human homologues, only two of which were previously known (1-B and 1-C). We propose to apply the existing nomenclature for the previously reported human exon 1s (Breslin et al. 2001) and extend the numbering for the new exons from D to H.

**3′ boundary locations of hGR exon 1s**

The 3′ boundary of each exon 1 with exon 2 was located by sequencing of cloned PCR products from total human cDNA. For each of the exon 1s, five positive colonies were sequenced. In all clones, a unique exon 1-specific splice donor site and a common exon 2 splice acceptor site were identified (Table 2). The human exon 2 splice acceptor site showed a remarkable homology with the corresponding rat sequence (human TTGATA, rat TTAATA). The common human exon 2 splice acceptor site is located 12 bp upstream of the ATG start codon, similar to the 14 bp observed in the rat. The exon 1 splice donor sites for exons 1-A and 1-B were confirmed. In addition, four new exon 1s have been identified and a fifth has been hypothesized on the basis of its homology to the rat. Following the existing human nomenclature, these were named 1-D to 1-H, and are homologous to the rat exons 14, 15, 17, 18 and 111 respectively. Interestingly, the region corresponding to the rat exon 19 appeared by simple PCR amplification.
to be spliced to the homologue of exon 1-10 (further examined below). Using RT-PCR, exon 1-G in Fig. 1, the human homologue of the rat exon 18 identified by alignment could not be detected in either universal human cDNA or the individual tissue samples tested.

5'-RACE identification of the exon 1 5' ends

To identify the start of the transcribed region for three of the four newly identified exon 1s (exons 1-E, 1-F and 1-H) cloned nested PCR products from the 5'-RACE reaction were sequenced. Sequencing of exon 1-E- and 1-F-specific colonies gave distinct exon 1 sequences of 47 and 62 bp respectively. These products, based on their determined 5' and 3' locations, showed an 80.0% and 69.8% homology with the corresponding exons 15 and 17 in the rat as shown in Fig 2.

Exon 1-E surprisingly produced a product that aligned almost perfectly with the rat exon 15 although it is 6 bp shorter. Exon 1-F contains two short inserts (11 bp and 4 bp) compared with its homologue 17 in the rat. Similarly exon 1-H, is almost perfectly aligned with its rat homologue 111, although some 18 bp longer.

The exact transcription start location of exon 1-D has not been accurately determined. 5'-RACE amplification of this exon was tried from the mRNA of a number of peripheral tissues, but no product was obtained. As shown below, this exon was found only in the hippocampus, for which no mRNA was available. Therefore the 5' terminus reported here was determined by PCR (data not shown) performed on commercially available cDNA with the forward primer location moved progressively towards the 5' transcription start site. The primer reported in Table 1 is the one with the longest positive product, and the 5' transcription start site reported in Table 2 represents the 5' location of this primer.

Exon 1-C has multiple splice variants

As mentioned above, the human homologue of rat exons 1-9 and 1-10 appears to be part of a single human exon, i.e. exon 1-C. Forward sequencing of the exon 1-9 and 1-10 homologues showed a common sequence of 17 bp at the 3' end, which is too short for placing an exon 1-C-specific RACE PCR primer. Therefore for exon 1C, the initial 5'-RACE PCR product (using the 2RACE RC primer) was cloned into the pCR4-TOPO vector. Colonies were screened using the 1-9-homologue primer, paired with the 2RACE RC reverse primer. Exon 1C positive clones were sequenced and produced three distinct products. These exon 1C splice variants shared both a common 3' terminus and a common 5' transcription initiation site. The 3' end of the human exon 1-C, as previously reported (Zong et al. 1990), has been sequenced as finishing 2371 bp upstream of the start of exon 2. Further examination of cloned exon 1-C 5'-RACE products confirmed that within exon 1-C there are at least three splice variants, represented in Fig. 3.
<table>
<thead>
<tr>
<th>Human exons</th>
<th>Rat exons</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-D</td>
<td>14</td>
</tr>
<tr>
<td>1-E</td>
<td>15</td>
</tr>
<tr>
<td>1-B</td>
<td>16</td>
</tr>
<tr>
<td>1-F</td>
<td>17</td>
</tr>
<tr>
<td>1-C1</td>
<td>19</td>
</tr>
<tr>
<td>1-C2</td>
<td>19</td>
</tr>
<tr>
<td>1-C3</td>
<td>19</td>
</tr>
<tr>
<td>1-H</td>
<td>11</td>
</tr>
</tbody>
</table>

**Figure 2:** Human and rat GR exon 1s and their location within the CpG island. Numbering is with respect to the translation start site within exon 2. The start of the human exon 2 is at -13 bp. Exon 2 (-1.6 kbp downstream) and the human exon 1-A (-31 kbp upstream) are shown. Light shading, published rat exons; dark shading, known and newly identified human alternative exon 1s; alternative splice donor (▲) and acceptor (arrows) sites within exon 1-C. TATA represents the potential TATA box for the expression of exon 1-E.
Exon 1-C appears itself to have three internal splice donor sites (SD1 to SD3) and three splice acceptor sites (SA1 to SA3) in each of two sub-exons, allowing a maximum of nine possible re-arrangements. However, using the 5'-RACE technique, only the three intra-exon splice variants shown in Fig. 3 (1-C1, 1-C2 and 1-C3) have been observed. Although named exon 1-C, this was the first hGR exon 1 reported.

**Tissue distribution of hGR exon 1s**

Human cDNA, reverse transcribed from blood cells isolated from normal volunteers and organ tissue biopsy samples was used to investigate expression patterns of the alternatively spliced exon 1s reported here. It is clear from Fig. 4 that these exons are differentially expressed in tissues. PCR primers located in the coding region of exon 2 detected the hGR in all tissues studied. However, as shown for the rat (McCormick et al. 2000), the human hippocampus was the only tissue expressing all the alternative splice variants. Within the limits of our tissue panel, exon 1-D was observed exclusively in the

---

**Figure 3** Arrangement of exon 1-C of the hGR gene. (A) Locations of the six exons within the human exon 1C and their three experimentally confirmed splice variants 1-C1, 1-C2 and 1-C3. The location of the exon 1-C as determined by Zong et al. (1990) is included for comparison. (B) The homologous rat exons 19 and 110 are shown to the same scale indicating their overlaps with the observed exon 1-C variants.

---

**Figure 4** Tissue distribution of the hGR alternative exon 1s. Products of the detection PCR of the different exon 1s are shown for the different tissues. Water and total human cDNA were used as negative and positive controls respectively.
hippocampus. All of the exons examined showed a distinct expression profile: exons 1-E and 1-F, in addition to the hippocampus, seem to be biased towards the immune system. Exon 1-E was observed in both CD8+ T cells and monocytes, exon 1-F was observed in CD14+ B lymphocytes and BCDA2+ peripheral blood dendritic cells, although a faint band of exon 1-E also appeared in the thyroid. Of the new exons, exon 1-H seems to have the largest tissue distribution including liver, lung, smooth muscle and immune cells (CD4+, CD8+, CD14+, CD19+), as well as the hippocampus.

The tissue distribution of the previously reported exon 1-B and the three new splice variants of 1-C were also investigated. Due to the splice variants within exon 1-C itself, three separate PCRs were necessary: one to detect 1-C3 only; another with the forward primer crossing the internal 1-C2 splice detecting only the 1-C2 variant; and a third primer pair detecting all three splice variants. Figure 4 shows that exon 1-C, in particular the 1-C3 variant, would appear to be widely expressed. The expression of exon 1-C2 was limited to the hippocampus, CD14+ monocytes, and CD19+ B lymphocytes. By inference from the 1-C3 and 1-C_Total PCRs, it would appear that 1-C1 is expressed in the skin and the lungs, although if it was expressed in any of the 1-C5-positive tissues it was not possible to detect these two splice variants separately.

**Discussion**

Analysis of the 5' end of the hGR revealed at least 11 splice variants in the UTR region of exon 1. These variants are based on seven exon 1s, four of which were so far unknown. As previously reported for the rat, the new exons 1s reported here lie within a CpG-rich island, a region with high rat–human sequence homologies. This 3 kbp CG-rich region, contains seven alternative rat exons 1s, and their eight homologous human splice variants, including exons 1-B and 1-C previously reported (Zong et al. 1990, Breslin et al. 2001). All of the exon 1 variants have a unique splice donor site and share a common splice acceptor site in exon 2. This is similar to the equivalent splice donor and conserved splice acceptor sites in the rat (McCormick et al. 2000). The mRNA splice variants have an in-frame TGA stop codon starting 12 bp before the ATG translation start site. Thus the predicted translated mRNA sequence from all of these variants are identical. Also the mouse shares significant homologies in this region with the rat, but the untranslated first exons have not been investigated to the same extent (McCormick et al. 2000).

To name the four new exon 1s observed, together with the ephemeric 1-G, we have chosen to extend the alphabetical nomenclature of exon 1s, from 1-D to 1-H (Fig. 1). As a result, the chronological order of their discovery is maintained, although this does not fully reflect the topographical order of the exons in the 5' UTR.

The four new exons show remarkable similarity to their rat homologues (Fig. 1). Exon 1-D is the longest of the new exons (174 bp), starting and finishing 17 and 36 bp upstream of the corresponding 5' and 3' ends of the rat exon 1_4 (Fig. 2). Exon 1-E is only 6 bp longer than its homologue 1_5. Exon 1-F contains two short inserts of 11 and 6 bp when compared with the rat exon 1_7. 1-H is 18 bp longer than the corresponding rat exon 1_11. The human exon 1-G, homologue of the minor rat exon 1_19, was not experimentally detectable, perhaps because of very low expression.

In addition to these new exons, we found that the human exon 1-C occurs as three distinct splice variants, which cover the region homologous to the rat exon 1_10 and partially overlap with the sequence homologous to the rat 1_9 (Fig. 3). These 5' RACE results differ from the original work of Zong et al. (1990) who observed one continuous exon of 185 bp with multiple transcription initiation sites, sharing the same TGCCAGAG 3' terminus. This discrepancy may in part be explained by the fact that the primer extension assay and RNAse protection assay originally used by Zong et al. (1990) did not permit sequencing or otherwise identification of any upstream splicing events.

The human exon 1-A is localised 31 kbp upstream (Breslin et al. 2001) of the translation start site in a region probably corresponding to the distal rat exon 1s (exon 1_1 to 1_3) located at >15 kbp upstream of the translation start site and largely outside the CpG island that was investigated here. This upstream region of distal promoters and exons in both the human and the rat warrant further studies of their own. All of the alternative hGR exons 1s presented here were found to be transcribed in human tissue. As shown for the rat, the human hippocampus expresses mRNA of all the novel exons reported here, together with those previously reported. On the other hand, hGR mRNA containing the previously published human exon 1-B and the 1-C3 reported here have the broadest tissue distribution. The expression of the other exon 1s reported here seems to be more tissue-specific. While exon 1-D is only in the hippocampus, exons 1-E and 1-F are also detected in the immune system, and exon 1-H additionally in the liver, lung and smooth muscle.

Multiple tissue-specific promoters have been reported for several members of the steroid hormone receptor family (Kastner et al. 1990, Kwak et al. 1993, Flouriot et al. 1998). Exon 1A of the hGR has also been associated with its own promoter known to be activated by glucocorticoids (Breslin et al. 2001). The existence of individual proximal promoter regions for each exon 1 could explain tissue-specific control of GR expression,
e.g. by cortisol in T lymphocytes and the hippocampus. In contrast to the hippocampus, (Sapolsky et al. 1984), T lymphocytes increased their GR expression in response to corticosteroids (Eisen et al. 1988, Ramdas et al. 1999). This effect may be explained by the differential usage of distinct exon 1 promoters between the hippocampus (all exon 1s expressed), CD8+ (not expressing 1-C2, 1-D and 1-F) and CD4+ T lymphocytes (not expressing 1-C2, 1-D, 1-E and 1-F).

With multiple first exons each with their own promoter region, control of hGR expression would appear in a new light. The hGR was described as a housekeeping or constitutively expressed gene with promoters that contain multiple GC boxes and no TATA or TATA-like box (Breslin & Vedeckis 1998). Promoters of housekeeping genes generally lack TATA boxes, and often use multiple transcription initiation sites. Careful examination of the regions between the new exons revealed a TATA box upstream of exon 1-E, available for classical TATA-dependent gene regulation. However, all other proximal promoter regions of the novel exons do not contain a TATA or TATA-like box, but rather Sp1 (Dyan & Tjian 1983) binding sites, which are normally linked to constitutive gene expression.

Our new exon 1s, with putative individual proximal promoter regions, suggest that the regulation of the hGR is more complicated than previously thought. The relationship between the usage of previously known promoters and the generated protein isoform (GRα or GRβ) has been investigated (Pedersen & Vedeckis 2003). GRα and GRβ protein isoforms are the result of alternative splicing of the two exon 9s, α and β. The physiological role of GRβ is still somewhat controversial. GRβ was initially reported as a dominant negative inhibitor of GRα (Oakley et al. 1996, 1999), an effect that has not always been confirmed (Orii et al. 2002). It has been recently reported that competition between GRα and GRβ for limited quantities of the co-activator molecule GRIP1 may explain the dominant-negative effect of GRβ (Charmandari et al. 2005). The authors could not show a link between the α and β isoforms and the three first exons previously known.

Overall GR levels are known to be autoregulated by glucocorticoids such as dexamethasone. Detailed investigation of the alternative first exon usage in the rat provided no evidence that these exons play a role in this autoregulation (Freeman et al. 2004). In these studies expression of exon 1 variants was observed to be combinatorial and uniformly regulated by dexamethasone. Thus, in humans, similar to the rat, the alternative first exons may be responsible only for tissue distribution but not for GR autoregulation.

Within exon 2 of the hGR there are two in-frame alternative ATG translation initiation sites, 27 bp apart. The use of these two sites produces two protein isoforms of both the GRα and GRβ. The N terminus of the GR-A isoform of both GRα and GRβ is nine amino acids longer than the GR-B isoform. In recent studies, an increase in the level of exon 1-A-3 correlated with an increase in the GR-B/GR-A ratio after dexamethasone stimulation, resulting in the fine-tuning of the transactivation potential of the resulting hGR complex (Pedersen et al. 2004). Indeed GR-B is twice as effective as GR-A in gene transactivation, but they are similar in their relative repressive abilities (Yudt & Cidlowski 2001). The addition of the four new exon 1s, together with the elucidation of the structure of exon 1-C multiplies the ways by which the expression of the GR-A and GR-B isoforms of both the GRα and GRβ can be controlled.

In summary, we have shown that the 5′ region of the hGR is more complex than has previously been described and that the multiple new exon 1s share important homologies with the rat GR gene structure. All the exon 1s investigated were shown to have a tissue-specific distribution. We suggest that each of these untranslated first exons has a distinct proximal promoter region, providing a mechanism for tissue-specific expression of the GR isoforms.

Acknowledgements
We are grateful to Wim Ammerlaan and Joé Weber for their technical assistance. We also thank Dirk Hellhammer for his support and initiatives within the Graduate School of Psychobiology of the University of Trier (Germany). The authors declare that there is no conflict of interest that would prejudice the impartiality of this scientific work.

Funding
We acknowledge the support of the Ministry of Health who provided the fellowship to J D T and the CRP Santé for assistance.

References


Breslin MB, Geng CD & Vedeckis WV 2001 Multiple promoters exist in the human GR gene, one of which is activated by glucocorticoids. Molecular Endocrinology 15, 1381–1395.
Charmandari E, Chrousos GP, Ichijo T, Bhattacharyya N, Vottero A, Souvatzoglou E & Kino T 2005 The human glucocorticoid receptor (hGR) beta isoform suppresses the transcriptional activity of hGRalpha by interfering with formation of active coactivator complexes. Molecular Endocrinology 19 32–44.


Pedersen KB & Vedeckis WV 2003 Quantification and glucocorticoid regulation of glucocorticoid receptor transcripts in two human leukemic cell lines. Biochemistry 42 10979–10990.

Pedersen KB, Geng CD & Vedeckis WV 2004 Three mechanisms are involved in glucocorticoid receptor autoregulation in a human T-lymphoblast cell line. Biochemistry 43 10851–10858.


Wei P & Vedeckis WV 1997 Regulation of the glucocorticoid receptor gene by the AP-1 transcription factor. Endocrine 7 303–310.


Received 19 May 2005
Accepted 13 June 2005
Made available online as an Accepted Preprint 23 June 2005