Cloning of the Human Erythropoietin Receptor Gene

By Constance Tom Noguchi, Kyung S. Bae, Kyung Chin, Yuko Wada, Alan N. Schechter, and W. David Hankins

We have isolated and characterized a genomic clone of the human erythropoietin (Epo) receptor from a placental genomic library using a cDNA probe for the murine Epo receptor. The coding region spans about 6.5 kb with seven intervening sequences ranging in size from 81 bp to 2.1 kb. A stretch of 123 purines is found in the 5' region from –456 to –578 upstream from the first codon and flanking the Alu repetitive sequences located further upstream. The human Epo receptor contains a palindromic sequence 5' of the translated region that consists of an almost perfect inverted repeat of 12 nucleotides (CAGCTGC/G/C)TCCG] centered about G at –92 from the first codon. An inverted SP1 binding site (CCGCCC) and an inverted GATA-1 binding site (TTATCT) are located at positions –151 and –179, respectively, and CACCC sequences are located at –585 and further upstream. No TATA or CAAT sequences are in this 5' flanking region. However, this region as far as –275 has a 72% GC content compared with an overall GC content of 56%. A 1-kb BamHI fragment of the human Epo receptor containing 700 bp of sequences 5' of the coding region was transcribed in an in vitro transcription assay; initiation of transcription appeared to be around 132 ± 5 just downstream from the inverted SP1 site at –151. T1 analysis of human Epo receptor messenger RNA also maps the site of transcription initiation to this region. Within 180 nucleotides 5’ to the first exon are three regions with 70% or greater homology with the murine Epo receptor. The study of this gene, including its similarities with the murine Epo receptor, should help elucidate aspects of the transcriptional and possible translational control of the Epo receptor in human erythroid cells and thus its role in signal transduction and erythroid differentiation.

This is a US government work. There are no restrictions on its use.

ERYTHROPOIETIN (Epo) is manufactured in the kidney and stimulates the proliferation and differentiation of erythroid cells. This 34,000 d glycoprotein is believed to be the primary regulator of erythropoiesis. The action of Epo is initiated by its binding to a specific cell surface receptor followed by receptor-mediated endocytosis. However, the molecular details regarding the manner in which Epo binding to its receptor stimulates erythropoiesis remains unknown.

Human erythroid progenitor cells exhibit both high- and low-affinity Epo binding sites. Cross-linking Epo to its receptor shows pepitides of two molecular weights [about 85,000 and 100,000 d] for the human Epo receptor, similar to that for the murine Epo receptor. In culture, human burst-forming units-erythroid (BFU-E) purified from peripheral blood exhibit much lower numbers of Epo receptors than colony-forming units-erythroid (CFU-E) and the numbers of Epo receptors increase as BFU-E mature. Analysis of bone marrow cells indicates that specific binding to Epo is greatest in pronormoblasts and declines during later erythroid cell maturation. Human erythroid cell lines such as HEL, K562, OCI1, JK-1, TF-1, KU-812, and RM10 exhibit one or two classes of receptors and the number of classes of receptors does not appear to be related to Epo responsiveness. Epo receptors have also been observed in nonerythroid cell lines such as the human myeloma cell line MM-S1.

Structural details, including the amino acid sequence of the Epo receptor, were recently derived from the cloning of the murine Epo receptor. The murine Epo receptor cDNA was isolated from a cDNA library constructed from uninduced murine erythroleukemia (MEL) cells that contained less than 1,000 sites per cell. Sequence analysis of the murine Epo receptor cDNA predicts a 507 amino acid polypeptide. The murine Epo receptor has a typical amino-terminus 24 amino acid signal sequence and single membrane-spanning domain. The predicted molecular weight is 55,000 d. When transfected into COS cells, the cDNA can be expressed and produces a high- and low-affinity receptor. The characterization of the murine cDNA Epo receptor facilitated the isolation of the full-length murine Epo receptor gene, as well as the analogous human Epo receptor cDNA.

The human Epo receptor cDNA isolated using the murine Epo receptor cDNA predicts a 508 amino acid polypeptide with 82% homology with the murine Epo receptor cDNA and contains an untranslated 3' tail about 200 bases long. This 3' untranslated region in the human Epo receptor cDNA exhibited unexpected homology with the 167-base untranslated 3' region in the murine cDNA clone.

To examine in detail the genomic sequences that may control transcription and expression of the human Epo receptor, we screened a human genomic library to isolate the Epo receptor gene. We identified a full-length genomic clone of 15 kb, sequenced the entire gene including 1.9 kb of 5' upstream sequence flanking the coding region, and examined the promoter activity in vitro of a fragment containing about 700 bases 5' of the coding region.

MATERIALS AND METHODS

A human placental DNA library constructed in cloning vector EMBL-3 SP6/T7 using a partial Sau3A digestion and the BamHI cloning site (Clontech Laboratories, Inc, Palo Alto, CA) was
screened for the Epo receptor. The library was grown on *Escherichia coli* strain NM538 with 10,000 to 30,000 plaque-forming units (pfu)/150-mm plate. The probe used for screening was the full-length murine Epo receptor cDNA contained in a 1,773-bp Kpn I fragment from pXM(ER)-190 labeled by nick translation with 32P-dCTP. Plaques (2.6 × 10^6) were screened after plaque transfer to nitrocellulose membranes by hybridization to probe in 50% formamide buffer incubated at 42°C for 20 hours. Secondary and tertiary screenings were used to isolate single positive clones.

To determine the size of the genomic DNA clone, restriction enzyme digestion with Sfi I was used to cleave the genomic insert with flanking T7 and SP6 promoters from the EMBL-3 SP6/T7 vector. For sequencing, the genomic clone was digested with *BamHI* and the resultant DNA fragments were subcloned into M13mp18 phage and pGEM7zf (+) plasmids (Promega Biotec, Madison, WI) as vectors for single-stranded and double-stranded sequencing, respectively. The largest fragment, 4.5 kb, was further digested with *PstI* and the generated fragments subcloned into pGEM5Zf (+) plasmid. Synthetic oligonucleotide sequencing primers were chosen based on the initial sequencing using primers from flanking vector sequences from M13 or from the SP6 and T7 promoters. Sequencing was performed in both the sense and antisense directions using the dideoxyoligonucleotide sequencing method.\(^8\) Sequenase T7 DNA Polymerase (United States Biochemical, Cleveland, OH) was used for the sequencing reactions with \([\alpha-\text{32P}]\) dATP. To resolve sequence ambiguities, Taq DNA polymerase (United States Biochemical) and/or \(\beta\)-deaza-dGTP was used. The specific orientation and sequential arrangements of the *BamHI* fragments were determined by polymerase chain reaction (PCR) analysis of the full-length genomic insert across the *BamHI* restriction enzyme sites.\(^8\)

Computer analysis of the sequencing gel was performed using SpeedReader (IntelliGenetics, Inc, Mountain View, CA) and analysis of sequencing data was performed using PC/GENE (IntelliGenetics) and MacVector (International Biotechnologies, Inc., New Haven, CT). Homology with published DNA sequences (Genbank; IntelliGenetics) was determined by computer search.\(^8\)

In vitro transcription was performed using K562 nuclear extract,\(^9\) HeLa nuclear extract,\(^10\) and a 175 mmol/L ammonium sulfate fraction of KS62 nuclear extract that is enriched for globin gene transcription activity in vitro.\(^11\) The 1.0-kb *BamHI* fragment containing about 700 bases of the coding region and exon 1 of the human Epo receptor was inserted into puc18, digested with restriction enzymes HindIII and EcoRI and used as a DNA template. Transcription reactions (50 µL) contain 10 to 15 µg of linearized DNA templates, 12 mmol/L HEPES (pH 7.9), 12% (vol/vol) glycerol, 0.3 mmol/L dithiothreitol (DTT), 0.12 mmol/L EDTA, 60 mmol/L KCl, 12 mmol/L MgCl_2, 600 µmol/L unlabeled triphosphates (ATP, CTP, UTP), 25 µmol/L GTP, 10 µCi of ([\alpha-\text{32P}]GTP, and 20 to 100 µg of nuclear extract proteins incubated for 60 minutes at 30°C. The RNA transcripts were resolved by electrophoresis on a 6% acrylamide-7 mol/L urea gel with the running buffer of Tris-borate-EDTA.

For T1 analysis,\(^25\) of the human Epo receptor messenger RNA (mRNA), total cellular RNA was prepared from OCI11 cells.\(^7\) A 793-bp *PstI* fragment was excised from the 1.0-kb *BamHI* fragment, subcloned into pGEM5Zf (+), and used to manufacture a \(\beta\)-labeled RNA probe using SP6 RNA polymerase. Ten micrograms of total RNA was hybridized with the \(\beta\)-labeled RNA probe and digested with ribonuclease T1 to excise regions that did not hybridize. The undigested RNA fragments were resolved by polyacrylamide gel electrophoresis.

**RESULTS**

We have isolated a 15-kb clone containing the Epo receptor gene (Fig 1A) from a human placental genomic DNA library using the full-length murine Epo receptor cDNA as probe. The genomic library was constructed using the EMBL-3 SP6/T7 cloning vector with flanking T7 and SP6 promoters and Sfi I sites. From 260,000 plaques of the human genomic library screened, 10 clones were identified to hybridize specifically to the murine Epo receptor cDNA probe. When cut with the restriction enzyme *BamHI*, these clones gave similar banding patterns when analyzed by agarose gel electrophoresis (Fig 1B). The major differences appeared with the high molecular weight fragments, indicating differences in the amount of flanking region and orientation of the insert in the cloning vector. Clone no. 26-6 was chosen for further characterization and sequencing. Digestion with the restriction enzyme Sfi I removes the flanking T7 and SP6 promoters along with the inserted genomic DNA fragment. Because no internal Sfi I restriction sites were observed within the isolated clones, we were able to determine the size of the insert at about 15 kb.

Digestion of the genomic clone with *BamHI* resulted in restriction fragments of 0.5, 1.0, 1.2, 1.6, and 4.5 kb and two long fragments (> 9 kb) corresponding to the remaining \(\lambda\) phase arms with attached flanking genomic sequences (Fig 1B). The five shorter *BamHI* fragments were subcloned and sequenced in both sense and antisense directions (Fig 2). Orientation and alignment of the *BamHI* fragments were determined by sequence homology to the murine cDNA. Confirmation of the orientation of the *BamHI* fragments, particularly the flanking 1.2-kb *BamHI* fragment was obtained by PCR analysis across the *BamHI* restriction enzyme site using the original genomic clone with subsequent cleavage of the resulting PCR fragment with *BamHI*. Such an analysis showed that no short fragments were lost or overlooked during the original subcloning. During the library construction, the original *BamHI* restriction site in the multiple cloning region of the vector was destroyed. Consequently, the *BamHI* digestion left adjacent genomic fragments attached to the arms of the \(\lambda\) phase vector. These flanking regions, which were 0.5 kb and 7 kb, respectively (determined by digestion with \(Xho\) I and \(Sac\) I, which excise the entire genomic insert), were not sequenced (Fig 1A).

The coding region for its 508 amino acids that is about 80% homologous to the murine Epo receptor cDNA is separated into 8 exons and distributed over about 7 kb of DNA (Figs 1B and 2). The borders between the exons and intervening sequences (IVS) (Table 1) were determined by comparisons between the genomic sequence and cDNA sequence generated from PCR cloning from OCI11 mRNA using primers determined from the genomic sequence (Hankins et al, unpublished data) and published cDNA sequence.\(^12\) The specific flanking donor and acceptor splice sites and the exon and IVS borders are shown in Table 1. The exons range in size and code for 29 (exons 6 and 7) to 203 (exon 8) amino acids. The coding region of the human Epo receptor was found to be identical to that reported by Jones et al.\(^13\) The intervening sequences range...
in size from 0.08 kb to 2.1 kb and contain stretches with a high degree of homology to Alu repeat sequences. These Alu repeat sequences are represented by off diagonal lines in the self homology matrix analysis of the human Epo receptor illustrated in Fig 3. The diagonal line represents the expected 100% self homology. As with the murine clone, exon 1 includes the 5' untranslated region, the signal peptide domain, and the amino-terminal portion of the extracellular domain. The entire extracellular domain is coded in 5 exons, the single transmembrane domain is coded in exon 6, and the cytoplasmic domain is localized on a very short exon 7 and a long exon 8. The size of the exons range from 87 (exons 6 and 7) to greater than 800 bp (exon 8). The gene does not contain a poly(A) consensus addition signal (AATAAA) in the 3' region.16

From the homology matrix comparison of the human Epo receptor with itself, the regions of Alu sequences are readily apparent (Fig 3). These regions include the 5' upstream region of the human Epo receptor (from −586 to −789 and −961 to −1915), IVS 2 (from 1420 to 1985), and IVS 6 (from 4205 to 4500, 4611 to 4895, and 5025 to 5130). The shorter intervening sequences (<900 bp) do not contain these repeat sequences. The GC content of the region 5' of the coding region extending to −275 is increased to 72% compared with an overall GC content of 56%, and this region presumably includes the promoter for the human Epo receptor. The region from −456 to −578 upstream from the ATG start site for translation consists of 123 purines including 13 repeats of GGAA and multiple repeats of GA. The murine Epo receptor sequence 5' of the coding region contains a purine rich region from −32 to −97. This GA-rich region (55 of 61 bases) contains a PU box (GAGGAA) that is the consensus sequence for the PU.117 or Sp1-1 DNA binding protein. In contrast, the human Epo receptor sequence contains a shorter GA-rich region (23 of 26 bases) immediately 5' of the coding region from −31 to −56. Unlike the murine sequence, neither this purine-rich region or the longer region further upstream in the human Epo receptor 5' region contains the PU.1 binding consensus sequence (GAGGAA).

Comparison of the 180 nucleotides in the 5' region flanking the coding region of the human Epo receptor shows three regions of homology from 70% to 89% with the murine sequence (Figs 1C and 4).16 The sequence immediately 5' of the coding region extending from −1 to −56 is 70% homologous with the comparable murine sequence. The purine-rich region extending from −32 upstream is longer in the murine 5' flanking region. Hence, the next region of homology is between −59 and −116 in the human sequence and −90 to −144 in the murine sequence. Specifically, the human sequences from −59 to −86 and from −91 to −116 are 89% and 88% homologous to the murine sequences from −90 to −117 and from −119 to −144, respectively. An overlapping region of homology exists between the human sequence from −117 to −185 and the murine sequence from −130 to −200. The human sequence from −117 to −145 is 86% homologous to the murine sequence from −130 to −158. This region of the
Fig 2. The 8.6 kb of the human Epo receptor gene. The coding region is underlined with the corresponding amino acids indicated.
Table 1. Exon and IVS Borders for Human Epo Receptor

<table>
<thead>
<tr>
<th>Exon</th>
<th>IVS 1</th>
<th>IVS 2</th>
<th>IVS 3</th>
<th>IVS 4</th>
<th>IVS 5</th>
<th>IVS 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>...ATG GAC CAC...</td>
<td>...AGC AAA G...</td>
<td>...ATG GAC CAC...</td>
<td>...AGC AAA G...</td>
<td>...ATG GAC CAC...</td>
<td>...AGC AAA G...</td>
</tr>
<tr>
<td>2</td>
<td>Met Asp His...</td>
<td>Ser Lys Ala...</td>
<td>Met Asp His...</td>
<td>Ser Lys Ala...</td>
<td>Met Asp His...</td>
<td>Ser Lys Ala...</td>
</tr>
<tr>
<td>3</td>
<td>...1 2 3...</td>
<td>...37 38 39...</td>
<td>...1 2 3...</td>
<td>...37 38 39...</td>
<td>...1 2 3...</td>
<td>...37 38 39...</td>
</tr>
<tr>
<td>4</td>
<td>972</td>
<td>1107</td>
<td>972</td>
<td>1107</td>
<td>972</td>
<td>1107</td>
</tr>
<tr>
<td>5</td>
<td>IVS 1: GCCCTCCCAG</td>
<td>CG GCC TTG...</td>
<td>Exon 2: CAG CTC GA GAG...</td>
<td>Exon 2: CAG CTC GA GAG...</td>
<td>Exon 2: CAG CTC GA GAG...</td>
<td>Exon 2: CAG CTC GA GAG...</td>
</tr>
<tr>
<td>6</td>
<td>Ala Ala Leu...</td>
<td>Gin Leu Glu...</td>
<td>Ala Ala Leu...</td>
<td>Gin Leu Glu...</td>
<td>Ala Ala Leu...</td>
<td>Gin Leu Glu...</td>
</tr>
<tr>
<td>7</td>
<td>39 40 41</td>
<td>82 83 84</td>
<td>39 40 41</td>
<td>82 83 84</td>
<td>39 40 41</td>
<td>82 83 84</td>
</tr>
<tr>
<td>8</td>
<td>2096</td>
<td>2270</td>
<td>2096</td>
<td>2270</td>
<td>2096</td>
<td>2270</td>
</tr>
<tr>
<td>10</td>
<td>Glu Asp...</td>
<td>Glu...</td>
<td>Glu Asp...</td>
<td>Glu...</td>
<td>Glu Asp...</td>
<td>Glu...</td>
</tr>
<tr>
<td>11</td>
<td>84 85 86</td>
<td>141 142 143</td>
<td>84 85 86</td>
<td>141 142 143</td>
<td>84 85 86</td>
<td>141 142 143</td>
</tr>
<tr>
<td>12</td>
<td>2351</td>
<td>2608</td>
<td>2351</td>
<td>2608</td>
<td>2351</td>
<td>2608</td>
</tr>
<tr>
<td>13</td>
<td>IVS 3: CCATGCCTGCA...</td>
<td>...GTA CAG AGG...</td>
<td>...GTA CAG AGG...</td>
<td>...GTA CAG AGG...</td>
<td>...GTA CAG AGG...</td>
<td>...GTA CAG AGG...</td>
</tr>
<tr>
<td>14</td>
<td>Ala Ala...</td>
<td>Leu...</td>
<td>Ala Ala...</td>
<td>Leu...</td>
<td>Ala Ala...</td>
<td>Leu...</td>
</tr>
<tr>
<td>15</td>
<td>143 144 145</td>
<td>193 194 195</td>
<td>143 144 145</td>
<td>193 194 195</td>
<td>143 144 145</td>
<td>193 194 195</td>
</tr>
<tr>
<td>16</td>
<td>2992</td>
<td>3145</td>
<td>2992</td>
<td>3145</td>
<td>2992</td>
<td>3145</td>
</tr>
<tr>
<td>17</td>
<td>IVS 4: TTCCGCCCATAG</td>
<td>GTG GAG ATC...</td>
<td>...Exon 5: CCT AGC G...</td>
<td>...Exon 5: CCT AGC G...</td>
<td>...Exon 5: CCT AGC G...</td>
<td>...Exon 5: CCT AGC G...</td>
</tr>
<tr>
<td>18</td>
<td>Glu Glu...</td>
<td>Leu...</td>
<td>Glu Glu...</td>
<td>Leu...</td>
<td>Glu Glu...</td>
<td>Leu...</td>
</tr>
<tr>
<td>19</td>
<td>196 197 198</td>
<td>246 246 247</td>
<td>196 197 198</td>
<td>246 246 247</td>
<td>196 197 198</td>
<td>246 246 247</td>
</tr>
<tr>
<td>20</td>
<td>3230</td>
<td>3371</td>
<td>3230</td>
<td>3371</td>
<td>3230</td>
<td>3371</td>
</tr>
<tr>
<td>21</td>
<td>IVS 5: CCGCCTTCTAG</td>
<td>AC CTG GAC...</td>
<td>...Exon 6: CAC CGC CG...</td>
<td>...Exon 6: CAC CGC CG...</td>
<td>...Exon 6: CAC CGC CG...</td>
<td>...Exon 6: CAC CGC CG...</td>
</tr>
<tr>
<td>22</td>
<td>Asp...</td>
<td>Arg...</td>
<td>Asp...</td>
<td>Arg...</td>
<td>Asp...</td>
<td>Arg...</td>
</tr>
<tr>
<td>23</td>
<td>247 248 249</td>
<td>274 275 276</td>
<td>247 248 249</td>
<td>274 275 276</td>
<td>247 248 249</td>
<td>274 275 276</td>
</tr>
<tr>
<td>24</td>
<td>5419</td>
<td>5606</td>
<td>5419</td>
<td>5606</td>
<td>5419</td>
<td>5606</td>
</tr>
<tr>
<td>25</td>
<td>IVS 6: TGGCTACACAG</td>
<td>G GCT GTC...</td>
<td>...Exon 7: AAC TTC CAG TAG...</td>
<td>...Exon 7: AAC TTC CAG TAG...</td>
<td>...Exon 7: AAC TTC CAG TAG...</td>
<td>...Exon 7: AAC TTC CAG TAG...</td>
</tr>
<tr>
<td>26</td>
<td>Arg...</td>
<td>Asp...</td>
<td>Arg...</td>
<td>Asp...</td>
<td>Arg...</td>
<td>Asp...</td>
</tr>
<tr>
<td>27</td>
<td>276 277 278</td>
<td>303 304 305</td>
<td>276 277 278</td>
<td>303 304 305</td>
<td>276 277 278</td>
<td>303 304 305</td>
</tr>
<tr>
<td>28</td>
<td>5602</td>
<td>6213</td>
<td>5602</td>
<td>6213</td>
<td>5602</td>
<td>6213</td>
</tr>
<tr>
<td>29</td>
<td>IVS 7: CTATTTCCTTAG</td>
<td>CTG TGG CTG...</td>
<td>...Exon 8: TGC TCT TAG GAC...</td>
<td>...Exon 8: TGC TCT TAG GAC...</td>
<td>...Exon 8: TGC TCT TAG GAC...</td>
<td>...Exon 8: TGC TCT TAG GAC...</td>
</tr>
<tr>
<td>30</td>
<td>Leu Trp...</td>
<td>Cys Ser...</td>
<td>Leu Trp...</td>
<td>Cys Ser...</td>
<td>Leu Trp...</td>
<td>Cys Ser...</td>
</tr>
<tr>
<td>31</td>
<td>306 307 308</td>
<td>507 508 509</td>
<td>306 307 308</td>
<td>507 508 509</td>
<td>306 307 308</td>
<td>507 508 509</td>
</tr>
</tbody>
</table>

Nucleotide numbers, from the ATG start codon, and amino acid numbers, from the first amino acid residue, are indicated for the beginning and end of each exon.

The human Epo receptor sequence contains the sites for initiation of transcription at -146 and -147, and to a lesser extent at -150 (which would be comparable to -133, -134, and -137 in the human sequence). However, other possible alternate start sites in the murine Epo receptor have been identified downstream from this region. The homology continues with the human sequence from -80 to -104 that is 80% homologous to the murine sequence from -59 to -80, and contains the possible regulatory sequences binding to SP1 (CCGCCCC)20 and GATA-1 (TTATCT)21,22 analogous to the sequences identified in the murine promoter region.

The human Epo receptor sequence from -80 to -104 contains a palindromic sequence CAGCTGCGTCCGGCGAGGCAGCTG that consists of an almost perfect inverted repeat of 12 nucleotides (CAGCTGC(G/C)TCCG) flanking a G at position -92. Although the region from -59 to -80 contains several regions of homology off the diagonal line indicating repeat sequences are detected in the 5' region and in IVS 1 and IVS 6. The dark cluster region on the diagonal line extending from 1338 to 1460 consists only of purines (from -456 to -578 5' of the coding region).
HUMAN ERYTHROPOIETIN RECEPTOR GENE

Fig 4. Homology matrix analysis of the 5' region flanking the first codon between the human and murine Epo receptor sequences. The search window is 25 nucleotides. Region A extends from −117 to −185 in the human and from −130 to −200 in the murine sequences. Region B extends from −59 to −116 in the human and from −90 to −144 in the murine sequences. Region C extends from −1 to −56 in the human and murine sequences. Sequence alignment is given in Fig 1.

to −116 in the human sequence is about 84% homologous with the murine sequence from −90 to −144, the murine sequence does not contain this inverted repeat and differs in 5 bases in the alignment with this palindromic region. However, the extensive homology surrounding this region suggests that this region may be important in transcriptional or translational control of the Epo receptor.

In the 5' region of the human Epo receptor gene the potential protein binding CACCC sequence found in the promoter region of globin genes32 is observed in the human Epo receptor at −585, −1705, and −1906 and the inverted GGGTG sequence appears at −50, −599, and −933. This regulatory sequence occurs three times within 481 bp 5' of the coding region in the murine Epo receptor gene.

As with the murine sequence, the human Epo receptor does not contain a TATA or CAAT sequence upstream from the cap site. The TATA and CAAT sequences in the vicinity (about 20 bp 5' for TATA and farther upstream for CAAT) of the start site for transcription generally provide binding sites for transcription factors and RNA polymerase II to support the “basal” level of transcription.33

We used an in vitro transcription system22,24 to examine the activity of the potential promoter region of the human Epo receptor and to determine the start site for transcription. The 1.0-kb BamHI fragment that contains about 700 bases 5' of the coding region, and presumably the human Epo receptor promoter, was inserted into puc 18 and used as a DNA template for in vitro transcription. Nuclear extracts from K562 cells that express low levels of Epo receptor and HeLa cells were used for transcription in vitro. HeLa cells do not express the Epo receptor (Hankins et al, unpublished observations) but provide a source for transcriptionally permissive nuclear extracts.24 The DNA template was sufficient to produce RNA transcripts in both K562 and Hela nuclear extracts (Fig 5). The size of the human Epo receptor transcript was 465 ± 5 nt, which includes 333 nt 3' of the translation start site extending 3' from AUG to the BamHI site and into the multiple cloning region of puc 18 and the HindIII site. This indicated that transcripts were initiated around 132 ± 5 nt 5' from the AUG translation start site. This is consistent with the transcription start site predicted from sequence comparison with the murine Epo receptor. The murine sequence contains sites for transcription initiation identified at −146 and −147, and to a lesser extent at −150.16 The homology with the murine sequence suggests that transcription for the human Epo receptor may initiate around −133 or −134 (corresponding to −146 or −147 in the murine sequence, respectively) (Fig 1C).

The transcription initiation site was also determined by T1 analysis25,26 of RNA from OCI1 cells7 that constitutively express the human Epo receptor. A 793-bp Pst I fragment from the human Epo receptor containing exon 1, a portion of IVS 1, and the 5' flanking region was used to construct the 32P-labeled probe. RNA fragments of 249 ± 1 and 241 ± 1 were detected after T1 RNase digestion (Fig 6). Taking into account the size of exon 1 (115 bp), the transcript size indicates that transcription initiation sites are located at −134 ± 1 and −126 ± 1. The longer transcript is in the region determined from in vitro transcription (Fig 5) and predicted from the homology with the murine sequence (Fig 1C).16

DISCUSSION

The human Epo receptor gene exhibits the general features associated with the murine Epo receptor4,37 and

Fig 5. In vitro transcription of the human Epo receptor. Lane 1 is the marker lane. Lanes 2, 3, and 4 are the transcripts generated using the human Epo receptor gene as DNA template and K562 nuclear extract, partial fractionation of K562 nuclear extract, and HeLa nuclear extract, respectively. Lanes 5 and 6 are transcripts using K562 extract and control DNA templates of puc18 and an ε-globin promoter plasmid, respectively. The size of the human Epo receptor transcript (ER) is 465 ± 5 nt. The high molecular weight plasmid transcripts (PT) are generated from transcription initiation sites on the parent plasmid. The size of the ε-globin promoter control transcript is 320 nt.
responsive murine erythroleukemia SKT6 cells, presumably resulting from alternate mRNA splicing. The corresponding truncated mRNA does not code for the transmembrane or cytoplasmic domains and results in a secreted form of the Epo receptor that may function to modify Epo response.

Several regions of homology exist between the human and murine Epo receptor sequences in the 5' coding region (Fig 4). Close alignment exists within the first 180 nucleotides with the exception of a purine-rich region consisting primarily of G and A nucleotides (55 of 61) between −32 and −97 in the murine sequence compared with a shorter region of G and A nucleotides (23 of 26) between −31 and −56 in the human sequence (Fig 1C). This purine-rich region is flanked on the 5' side by a region with 84% homology between the human (from −59 to −116) and murine (from −90 and −144) sequences. In this region the human sequence contains a long palindromic sequence consisting of an inverted repeat of 12 nucleotides flanking the G at −92 with only one mismatch. The murine sequence does not contain the inverted repeat and differs in 5 nucleotides in the alignment with this palindromic region.

In the human sequence, the palindromic region is 3' of the start sites for transcription suggested by the homology with the murine sequence and the possibility of secondary structure formation may explain the difficulty in obtaining accurate data on the initiation site for transcription. Based on the homology between the human sequence from −117 to −185 and the murine sequence from −130 to −200, the transcription start site around −147 in the murine sequence would correspond to about −134 in the human sequence. The region farther upstream contains several potential regulatory sequences, including an SPI binding site (CCGCC) at −151 for the human and −166 for the murine sequence respectively and an inverted GATA1 binding site at −179 for the human and −194 for the murine sequence, respectively. In contrast to the murine Epo receptor, which contains three CACCC sequences in the 5' flanking region at −261, −290, and −477, the human Epo receptor contains CACCC sequences farther upstream at −574, −1705, and −1906.

To determine if the isolated human Epo receptor gene was transcriptionally active, the gene was used in an in vitro transcription assay. Activity of the murine Epo receptor gene has previously been shown by the ability of the promoter region (−581 to −130) 5' from the coding region to direct transcription of a reporter gene in MEL cells. We assumed that the promoter would be contained in a fragment of about 0.7 kb 5' upstream of the translation start site (AUG). This fragment was subcloned into a plasmid vector and was used as a DNA template for in vitro transcription driven by nuclear extracts from HeLa and K562 human erythroleukemia cells. We found that this fragment was transcriptionally active and mapped the initiation site of transcription to a region 132 ± 5 bp upstream of the translation start site (Fig 5). This is consistent with a cap site at about −134 that we predicted based on homology with the 5' murine sequence and is 3' of the potential SPI binding site at −143 to −148 and the inverted GATA-1 binding site at −170 to −175 (Fig 1C).
The transcription initiation site was also determined for the murine Epo receptor at approximately -137. The shorter transcript from the human Epo receptor at -126 ± 1 detected in the T1 protection assay was not observed in the in vitro transcription assay (Fig 5).

The amino acid sequence for the Epo receptor exhibits significant homology to other hematopoietic cytokine receptors. This superfamily of receptors includes the Epo receptor and receptors for interleukin-2 (IL-2), interferon-γ, tumor necrosis factor-α, and granulocyte-macrophage colony-stimulating factor (CSF), as well as the prolactin receptor. This supergene family of transmembrane receptors is characterized by one or more external domains that contain two conserved motifs separated by ~100 amino acids: a motif of about 60 amino acids containing four conserved cysteines that form disulfide bonds and a motif of about 20 amino acids located close to the transmembrane domain containing a Trp-Ser-X-Trp-Ser motif or the “WSXWS” box. The importance of the cysteine residues in the extracellular domain and their role in secondary structure is underscored by the development of a mutant murine Epo receptor with a substitution of a cysteine for an arginine at position 129 in the extracellular domain that is able to induce hormone-independent cell growth in IL-3-dependent hematopoietic Ba/F3 cells and render these cells tumorigenic in syngeneic mice. Although tyrosine phosphorylation of the receptor may occur after binding to the appropriate cytokine, the consensus sequence for tyrosine kinase does not appear in the cytoplasmic domain. The homology among the cytokine receptors in this super family of genes suggests possible similarities in their signal transduction pathway and their common motifs may participate in protein-protein interactions.

Although Epo is the principal regulator of erythropoiesis, sensitivity to Epo by erythroid progenitors appears to be developmentally controlled. For example, the early BFU-E that is capable of forming large bursts of erythroid cells is less sensitive to Epo than its more mature CFU-E counterpart that gives rise to smaller erythroid colonies. Detailed analysis of promoter elements and other cis-acting sequences may provide information regarding possible trans-acting factors that regulate the developmental expression of the Epo receptor and more generally the developmental response to Epo. Cloning of the Epo receptor provides a means for examining the developmental or differential regulation of this gene and possibly the early events of erythropoiesis, including before and after Epo stimulation.

ACKNOWLEDGMENT

The authors thank Dr Alan D’Andrea for the murine cDNA and Drs Hemin Chin, Donna Williams, and Feng-Sheng Xu for helpful discussions.

REFERENCES

7. Fraser JK, Lin FK, Berridge MV: Expression of high affinity receptors for erythropoietin on human bone marrow cells and on the human erythroleukemic cell line. HEMATOL 16:836, 1988
Cloning of the human erythropoietin receptor gene

CT Noguchi, KS Bae, K Chin, Y Wada, AN Schechter and WD Hankins

Updated information and services can be found at:
http://www.bloodjournal.org/content/78/10/2548.full.html

Articles on similar topics can be found in the following Blood collections

Information about reproducing this article in parts or in its entirety may be found online at:
http://www.bloodjournal.org/site/misc/rights.xhtml#repub_requests

Information about ordering reprints may be found online at:
http://www.bloodjournal.org/site/misc/rights.xhtml#reprints

Information about subscriptions and ASH membership may be found online at:
http://www.bloodjournal.org/site/subscriptions/index.xhtml