Genomic Structure, Chromosomal Localization, and Conserved Alternative Splice Forms of Thrombopoietin

By Austin L. Gurney, Wun-Jing Kuang, Ming-Hong Xie, Beth E. Malloy, Dan L. Eaton, and Frederic J. de Sauvage

Thrombopoietin (TPO), the ligand for c-mpl, is a novel cytokine comprising an amino terminal domain with homology to erythropoietin and a glycosylated carboxyl terminal domain that does not bear overall homology to other known proteins. We report the cloning of cDNAs encoding the porcine and murine TPO and the characterization of the human TPO gene. The cDNA for an additional splice form (TPO-2) with a four-amino-acid deletion within the erythropoietin-like domain has been isolated and is conserved between humans, pigs, and mice. Species comparison of TPO shows that the amino terminal erythropoietin-like domain is highly conserved, while the carboxyl terminal domain is less conserved. Recombinant murine TPO and human TPO are each able to activate both the murine and human c-mpl receptors, indicating an absence of strict species specificity. Human TPO is encoded by a single gene consisting of six exons and located on chromosome 3q27-28.

© 1995 by The American Society of Hematology.

MATERIALS AND METHODS

Isolation of the porcine TPO gene and cDNA. Genomic clones of the porcine TPO gene were isolated by screening a pig genomic library (EMBL3; Clontech Inc, Palo Alto, CA) with pR45, a previously described oligonucleotide probe. This probe was derived from a polymerase chain reaction (PCR) product obtained using amino acid sequence information from purified porcine TPO. Clones (2 × 10⁵) were lifted on nitrocellulose filters and hybridized overnight to pR45 in 35% formamide, 5× SSC, 50 mmol/L Na PO, pH 6.5, 0.1% sodium pyrophosphate, 10× Denhardt’s solution, 0.1% sodium dodecyl sulfate (SDS), 0.1 mg/ml salmon sperm DNA and washed for 45 minutes in 0.3× SSC, 0.1% SDS at 42°C. Several clones were isolated, and the exon corresponding to the amino acid sequence obtained from the purified porcine TPO was sequenced. The porcine TPO cDNA was obtained by rapid amplification of cDNA ends (RACE) PCR with the following modifications. Polyadenylated mRNA was isolated from the kidney of an aplastic pig, as previously described. Single-strand cDNA was first prepared by reverse transcription using BamR1: 5′GACTGAGGATCCATGATTTTTTTTTTTTTTTTTTTTTT 3′. An initial round of PCR amplification (28 cycles of 60 seconds at 95°C, 60 seconds at 58°C, and 90 seconds at 72°C) was conducted using the TPO-specific primer F-1 (5′GCTAGCTCTAGAATTGCTCCTCGTGGTCATGCTTCT 3′) and the shortened antisense primer BAMA (5′GACTGAGGATCCATC 3′) in a 100-μL reaction (50 mmol/L KCI, 1.5 mmol/L MgCl₂, 10 mmol/L Tris pH 8.0, 0.2 mmol/L dNTPs, with 0.05 U/μL AmpliTaq polymerase, Perkin Elmer Inc, Norwalk, CT). The PCR product was digested with Cla I and ligated to 0.1 μg of Bluescript SK vector (Stratagene Inc, La Jolla, CA), previously digested with Cla I and Kpn I. After incubation for 2 hours at room temperature, one fourth of the ligation mixture was subjected to a second round of PCR (22 cycles as described above) using the internal TPO-specific primer F-2 (5′GCTAGCTCTAGAAGC-CCCCGCTCTCTGCTGCTG 3′) and the TPO-specific primer F-3 (5′GCTAGCTCTAGAAGC-CCCCGCTCTCTGCTGCTG 3′) and the shortened antisense primer BAMA (5′GACTGAGGATCCATC 3′). The resulting PCR product was digested with Xba I and Cla I and subcloned into Bluescript SK−. Several clones from independent PCRs were sequenced. The DNA

From the Departments of Molecular Biology and Cardiovascular Research, Genentech, Inc, San Francisco, CA.

Address reprint requests to Frederic J. de Sauvage, PhD, Department of Molecular Biology, Genentech, Inc, 4600 Otis, San Bruno Blvd, San Francisco, CA 94080.

The publication costs of this article were defrayed in part by charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. section 1734 solely to indicate this fact.

© 1995 by The American Society of Hematology.

0006-4971/95/8504-0018$3.00/0
sequence of the first seven deduced amino acids of the mature TPO were derived from sequence contained within the PCR primers.

**Isolation of murine TPO cDNA.** A λgt10 cDNA library prepared from murine liver mRNA (Clontech, Inc) was screened with a probe encompassing the entire coding region of the human TPO cDNA. Clones (2 × 10^6) were screened essentially as described above with a probe labeled by the random oligonucleotide method using [32P]-labeled d-adenosine triphosphate (dATP) and [35P]-labeled d-cytidine triphosphate (dCTP). cDNA from purified phage were subcloned into the EcoRI site of Bluescript SK+ (Stratagene), and the DNA sequence of both strands was determined on an ABI373 automated sequencer.

**Isolation of the human TPO gene.** Human genomic DNA clones of the TPO gene were isolated from a human genomic library in λGem12, as previously described.

**Northern blot analysis.** A northern blot of poly (A)+ RNA from various mouse tissues (2 μg per lane) was purchased from Clontech and hybridized overnight with a random oligonucleotide-labeled probe containing the entire murine TPO coding region. The blot was washed at 42°C in 0.2× SSC/0.1% SDS and exposed overnight to X-omat film (Kodak, Rochester, NY).

**Expression of recombinant TPO.** Expression vectors for murine and human TPO were prepared as previously described. Clones encoding TPO and TPO-2 were subcloned into pRK5tkneo, a mammalian expression vector containing the cytomegalovirus (CMV) promoter and the SV40 polyadenylation signal. The resulting expression vectors (PRKtkneo-mTPO, PRKtkneo-mTPO-2, PRKtkneo-hTPO, and PRKtkneo-hTPO-2) were transiently transfected into 293 cells (a human embryonic kidney cell line) using the calcium phosphate method. After transient transfection, medium was conditioned for 3 days. Cells were maintained in high glucose Dulbecco’s modified Eagle’s medium (DMEM) supplemented with 10% fetal calf serum (FCS).

**Proliferation assays.** Stable cell lines expressing human or murine c-mpl in Ba/F3 cells were prepared as described previously. Ba/F3 cells were maintained in RPMI 1640 media supplemented with 10% FCS and 5% conditioned medium from WEHI-3B cells as a source of IL-3. For proliferation assays, cells were cultured in the absence of IL-3 for 16 hours in RPMI, 10% FCS. Cells were washed twice with PBS and plated in 96-well plates at 25,000 cells per well in the presence of various dilutions of conditioned medium from 293 cells transiently transfected with expression vectors encoding human or murine TPO. Each concentration was tested in duplicate. After 22 hours at 37°C, 1 μCi of [3H]-thymidine was added per well, and incubation was continued for an additional 6 hours. Incorporation of [3H]-thymidine was measured with a Top Count Counter (Packard Instruments).

**Immunoprecipitation of TPO.** Expression vectors encoding hu-

---

**Fig 1.** Comparison of the predicted mature TPO from mouse, pig, and human. The predicted amino acid sequence of murine, porcine, and human TPO (m-TPO, p-TPO, and h-TPO, respectively) are aligned. Identical amino acids are boxed, and gaps introduced for optimal alignment are indicated by dashes. Potential N-linked glycosylation sites are indicated by the shaded boxes. The conserved cysteine residues are indicated by dots. The conserved dibasic amino acid motif that presents a potential protease cleavage point is underlined. The four-amino-acid deletion found to occur in each species (TPO-2) is outlined by a box.
Porcine TPO cDNA was isolated by PCR using an oligonucleotide dT primer and two specific primers based on the genomic sequence of the exon of the TPO gene encoding the N-terminal amino acid sequence obtained from purified mature TPO. The cDNA was synthesized from various tissues of aplastic pigs and amplified using a modified RACE-PCR technique. A product of 1,342 bp was obtained from kidney mRNA and subcloned. Several clones were sequenced and found to encode the complete mature pig TPO. The cDNA encodes a 332-amino-acid mature protein (Fig 1). TPO-2, a second form, encoding a protein with a four-amino-acid deletion corresponding to amino acid residues 112 to 115, was also identified.

Murine TPO cDNAs were isolated from a mouse liver cDNA library using probes derived from the human TPO cDNA. The longest murine TPO cDNA clone contained an open reading frame of 1,068 bp flanked by 137 bases of 5' and 244 bases of 3' untranslated sequence followed by a poly(A) tail. The presumed initiation codon at nucleotide position 138 to 140 lies within a sequence context favorable for eukaryotic translation initiation. The open reading frame encodes a predicted protein of 356 amino acids. The amino terminus is highly hydrophobic and likely to function as a secretory signal, which is presumably cleaved between amino acids 21 and 22 to generate a mature protein of 335 amino acids.

Fig2. Northern blot analysis of murine TPO expression in selected tissues. Northern blot analysis of poly (A)^+ mRNA (2 µg per lane) isolated from the indicated tissues was hybridized overnight to a probe containing sequence from the entire murine TPO cDNA. The position of molecular weight markers is indicated.

Fig3. Determination of relative expression levels of TPO and TPO-2. The relative abundance of TPO and TPO-2 in the medium of transiently transfected 293 cells was determined by immunoprecipitation of TPO with anti-TPO antisera from [35S]-labeled conditioned medium. The 293 cells were transiently transfected with expression vectors for human TPO or TPO-2 or a control pRK5 vector. Immunoprecipitated proteins were separated by SDS-PAGE on a 4% to 12% gel and visualized by autoradiography. Lane 1 (prebleed) shows a control immunoprecipitation with preimmune sera of conditioned media containing human TPO. The mobility of molecular weight standards is indicated.
Effect of recombinant TPO on Ba/F3-mpl cell proliferation. (A) The ability of conditioned medium from 293 cells transiently transfected with expression vectors for murine TPO or TPO-2, human TPO, or control vector pRK to induce proliferation in Ba/F3 cells expressing murine c-mpl was measured by extent of incorporation of [3H]-thymidine into DNA. The average of duplicate determinations is shown. The experiment was repeated several times with similar results. (B) The ability of conditioned medium from 293 cells transiently transfected with expression vectors for human TPO or TPO-2, murine TPO, or control vector pRK to induce proliferation in Ba/F3 cells expressing human c-mpl was measured as in panel A.
Fig 5. Sequences of the human TPO gene. (A) A schematic of human TPO gene is shown. Exons are boxed. The hatched box indicates 5’ and 3’ untranslated sequence as identified in human TPO cDNA. The dark shaded boxes indicate the secretion signal. The light shaded boxes indicate sequence encoding the qm-like domain, and the white box indicates sequence encoding the carboxyl domain. The four amino acids involved in the alternative splice variants of TPO are located at the 5’ junction of exon 6 and are indicated by a black box. The ATG initiation codon and the stop codon are indicated. (B) The sequence of the TPO gene is shown in the 5’ to 3’ orientation. Exon coding sequence is denoted in upper case letters, and intron sequence is denoted in lower case letters. GT/AG splice donor and acceptor sites are indicated in bold characters. The deduced amino acid sequence of TPO is shown above the DNA sequence. The putative signal peptide is underlined. Cysteine residues are indicated by dots. The sequence involved in the alternative splice variants of TPO is double-underlined. A single base difference, indicated by vertical arrowhead, between the gene sequence and the reported cDNA sequence is due to a polymorphism and does not alter the amino acid sequence. The ATG translation start and the in-frame stop codon are boxed. The polyadenylation site is noted.
rine and human TPO cDNAs encode functional c-mpl ligands. However, comparison of the titration curves obtained in parallel experiments using conditioned medium containing human TPO versus murine TPO shows a degree of species specificity. Conditioned medium from cells transfected with expression vectors encoding either murine or human TPO-2 did not induce a proliferative response in Ba/F3-mpl cells. The ability to detect TPO activity at dilutions greater than the 500-fold difference in expression between TPO and TPO-2 suggests that the limited amount of TPO-2 that is expressed is not active.

Analysis of both human and murine genomic DNA by Southern blot indicates that the gene for TPO is present in a single copy (data not shown). The human TPO gene was isolated in two overlapping lambda clones spanning 35 kb. Two overlapping fragments (BamH1 and EcoRI) containing the entire TPO gene were subcloned and sequenced. The structure of the human gene is composed of six exons within 7 kb of genomic DNA (Fig 5). The boundaries of all exon/intron junctions are consistent with the consensus motif established for mammalian genes.17 Exon 1 and exon 2 contain 5' untranslated sequence and the initial four amino acids of the signal peptide. The remainder of the secretory signal and the first 26 amino acids of the mature protein are encoded within exon 3. The entire carboxyl domain and 3' untranslated, as well as approximately 50 amino acids of the epo-like domain, are encoded within exon 6. The four amino acids involved in the deletion observed within TPO-2 are encoded at the 5' end of exon 6.

The chromosomal location of the human TPO gene was determined by fluorescent in situ hybridization (FISH). The full-length human TPO cDNA was biotinylated and hybridized to the metaphase spreads of a normal individual. The 1.8-kb cDNA gave a clear FISH signal that mapped to 3q27-28 (Fig 6). The regional assignment was determined by the analysis of 20-well spread metaphases. A detectable hybridization signal at 3q27-28 was noted in 45 of 100 chromatid targets.

DISCUSSION

The cDNA of porcine TPO and murine TPO encode putative mature proteins of 332 amino acids and 335 amino acids, with molecular weights of 35.1 kD and 37.8 kD, respectively. TPO contains two distinct domains roughly corresponding to the amino- and carboxyl-terminal halves of the protein. The amino-terminal 153 amino acids of TPO bear homology to epo (23% identity, 50% similarity, including conservative substitutions with human TPO). Three of the four cysteines present within epo, including the two found to be essential for epo activity,18 are conserved in TPO, suggesting that this domain may assume a four α-helix bundle configuration similar to other members of the hematopoietic cytokine family.19
The carboxyl domain of TPO is not present in epo or any of the other hematopoietic growth factors identified to date. Interestingly, all the potential N-linked glycosylation sites of TPO from each species are confined to the carboxyl domain of the protein. Immunoprecipitation of human TPO shows a protein of 68 to 85 kD, suggesting that this domain is highly glycosylated as the predicted molecular weight for the nonglycosylated molecule is 38 kD. The function of this highly glycosylated domain remains to be determined. Glycosylation of the carboxyl domain of TPO may provide a longer half-life for TPO in vivo, as in the case for epo.19,20

TPO does not display strict species specificity. Porcine TPO was purified based on its ability to bind to recombinant human mpl-IgG fusion protein and stimulate [3H]-thymidine incorporation in Ba/F3 cells expressing human c-mpl, and recombinant human TPO induced platelet production in a mouse rebound thrombocytosis assay.5 Here we demonstrate that recombinant murine TPO and human TPO are able to induce a proliferative response in Ba/F3 cells transfected with either murine c-mpl or human c-mpl. The lack of strict species specificity is consistent with the high degree of conservation within the amino-terminal domain. There is, however, a degree of species specificity. Human and murine TPO expression in the conditioned medium of transiently transfected 293 cells was similar as estimated by immunoprecipitation using polyclonal antisera raised against the human TPO epo-like domain (data not shown). This suggests that the human ligand is less active than the murine ligand on the murine receptor.

Two alternative splice forms of TPO have been identified. They differ by the presence (TPO) or absence (TPO-2) of a four-amino-acid insertion at positions 112 to 115. The insertion/deletion occurs within the amino terminal epo-like domain at an identical position within murine, porcine, and human TPO. The four amino acids are encoded at the 5’ junction of exon 6 in the human gene. The TPO-2 mRNA form may result from the use of a favorable alternative splice acceptor sequence. To date, no splice variants of epo have been identified. However, granulocyte colony-stimulating factor (G-CSF) does possess two alternative splice variants that differ by a three-amino-acid deletion/insertion. Both of these forms are active, although the larger splice form is reported to have a lower specific activity.23 In contrast, we have not observed any agonist activity with TPO-2, the shorter form. Immunoprecipitation experiments suggest that TPO-2 is very poorly secreted. A similar deletion has been introduced into epo to study epo function.24 This five-amino-acid deletion of residues 122 to 126 alters a loop between alpha helices C and D in the proposed structure of epo and results in an inhibition of secretory. By comparison with epo, the four-amino-acid deletion within TPO-2 would also be predicted to occur within this loop. Thus, the significance of TPO-2 is uncertain. As mRNA encoding the shorter TPO form appears to be abundant in each of the three species examined to date, this may represent a novel mechanism for cytokine regulation. The level of TPO expression could potentially be modulated by altering the proportion of TPO mRNA that is spliced to encode each of the forms.

The human TPO gene was mapped by FISH to chromo-

some 3q27-28. Although no other cytokine gene has been localized to this locus, it will be of interest to examine this region, as there are several examples in the human and the mouse genome of close tandem linkage of genes encoding growth factors or growth factor receptors. The first to be described was a region of chromosome 5 containing the genes coding for granulocyte-macrophage (GM)-CSF, IL-3, IL-4, and IL-5.25,26 The mapping of the TPO gene is particularly interesting, as structural abnormalities of the long arm of chromosome 3 have been associated with increased bone marrow megakaryocytes and elevated platelet counts in patients with acute nonlymphocytic leukemia.10,12 Although frequently found in combination with other chromosomal defects, the most common abnormality in these patients is an inversion of 3q between bands 21 and 26,2 suggesting that a gene critical to the regulation of thrombopoiesis may be located at one end of the inverted segment. In fact, elevated levels of TPO activity have been measured in the serum of patients with 3q inversion.27 Characterization of the TPO gene and of its promoter in these patients will help to determine if aberrant production of TPO is responsible for this leukemia.

ACKNOWLEDGMENT

We thank Audrey Goddard, Jeremy Squire, and Barbara Beatty for assistance in FISH analysis; Greg Bennett for antibodies; the Genentech DNA synthesis group; L. Tamayo and A. Bruce for preparing figures; Sunita Sohrabji for help in preparing the manuscript; and Michel Aguet for criticism of the manuscript.

REFERENCES


From www.bloodjournal.org by guest on January 10, 2018. For personal use only.


Genomic structure, chromosomal localization, and conserved alternative splice forms of thrombopoietin

AL Gurney, WJ Kuang, MH Xie, BE Malloy, DL Eaton and FJ de Sauvage