Complex Pattern of Alternative Splicing in the Normal Uroporphyrinogen Decarboxylase Gene: Implications for Diagnosis of Familial Porphyria Cutanea Tarda

Julie F. McManus, C. Glenn Begley, and Sujiva Ratnaike

We describe multiple alternative transcripts of uroporphyrinogen decarboxylase mRNA in normal individuals and patients with familial porphyria cutanea tarda. mRNA was reverse-transcribed, subjected to the polymerase chain reaction, and analyzed for nucleotide sequence. Seven different transcripts were characterized, and a cryptic splice acceptor site was identified in intron 1. In all mRNAs the exons abutted at previously defined exon boundaries. Characterization of the splice junctions in the genomic DNA showed that splice donor and acceptor sequences complied with the consensus sequences for these sites except for the splice acceptor sequences of exons 3 and 10. These deviations were present in two normal individuals and one patient with familial porphyria cutanea tarda and were thus unable to explain the multiple aberrant uroporphyrinogen decarboxylase transcripts. We conclude that apparent deletions observed in transcripts derived from the uroporphyrinogen decarboxylase gene in patients with familial porphyria cutanea tarda should be interpreted with caution.

Indexing Terms: polymerase chain reaction/splice junctions/nucleotide sequencing.

The gene for human uroporphyrinogen decarboxylase (UROD; EC 4.1.1.37) is a housekeeping gene that codes for the fifth enzyme in the heme biosynthetic pathway. This enzyme decarboxylates the eight-carboxylic acid porphyrinogens, uroporphyrinogen I and III, to the four-carboxylic acid porphyrinogens, coproporphyrinogen I and III (4). The order of decarboxylation of the carboxylic acid residues is controversial (2, 3) and the catalytic site(s) unknown.

The UROD gene has been assigned to chromosome 1 (4). Both the gene for UROD and its cDNA have been cloned and sequenced and comprise 10 exons spread over 8 kilobases (kb) (5, 6). Abnormalities in the UROD gene have also been shown to be causative in familial porphyria cutanea tarda (fPCT) (7) and hepatoerythropoietic porphyria (8–10). In fPCT the UROD activity in liver and erythrocytes is ~50% of normal, whereas in hepatoerythropoietic porphyria the UROD activity is <10% (11).

Several studies have examined the molecular basis for fPCT. A point mutation has been described in the UROD gene cDNA in one family with fPCT. In this family, a G→C substitution (GGG→GCG) at position 860 of the cDNA resulted in valine replacing glycine at amino acid position 281 and resulted in a catalytically active but unstable protein (7). More recently, Garey et al. (12) described a deletion of exon 6 from the mRNA, which was attributed to a G→C substitution at the 5' end of intron 6. This defect was observed in 5 of 22 unrelated families with fPCT. However, in a number of other families, no defect in the UROD gene cDNA was defined.

In this study, we examined two normal individuals, an index patient with fPCT, and one of the index patient's offspring, also diagnosed with fPCT. We observed multiple UROD transcripts in all four subjects and, in all of them except the proband's daughter, we characterized the intron/exon boundaries of their UROD mRNA.

Materials and Methods

Subjects

The two normal individuals, #1 and #2, were healthy laboratory workers without any clinical or biochemical evidence of porphyria cutanea tarda (PCT). The index patient and her daughter were diagnosed with fPCT on the basis of skin photosensitivity typically associated with PCT and confirmatory porphyrin biochemistry. Concentrations of porphyrins in their urine, feces, and plasma were increased, and isocoproporphyrin, a diagnostic marker of PCT (13), was detected in their feces. Erythrocyte UROD activity was decreased in the proband as well as in three other family members with clinical PCT, including the aforementioned daughter. The procedure for handling the subjects used in this study was in accordance with the guidelines of the Royal Melbourne Hospital ethics committee.

Procedures

Erythrocyte UROD activity measurement. UROD activity was measured in the erythrocytes of both normal individuals and patients with fPCT. A sample of whole blood was incubated with pentacarboxylic acid porphyrinogen 1 (oxidized form; Porphyrin Products, Logan, UT) at 37°C for 30 min at pH 6. The reaction was terminated by the addition of an equivalent mixture of 0.61 mmol/L trichloroacetic acid and dimethyl sulfoxide containing mesoporphyrin (Porphyrin Products) as an internal standard. The coproporphyrin produced was
detected directly by HPLC with fluorescence detection and compared with a coproporphyrin III (Porphyrin Products) standard (14).

RNA isolation, cDNA synthesis, and extraction of genomic DNA. Mononuclear cells were isolated from 40 mL of peripheral blood by a Lymphoprep (Nycomed Pharma AS, Oslo, Norway) gradient, and the buffy coat was removed from centrifuged whole blood, and the leukocytes were washed in 8 g/L ammonium chloride for 15 min at 37°C. Total RNA was extracted from the leukocytes with acid guanidinium thiocyanate (International Biotechnologies, New Haven, CT) (15). cDNA was synthesized from at least 25 μg of total RNA by using avian myeloblastosis virus reverse transcriptase according to the manufacturer's instructions (Boehringer Mannheim, Mannheim, Germany). Poly(A)+ mRNA for Northern analysis was extracted with oligo dT cellulose (Pharmacia, Uppsala, Sweden) by standard techniques. Genomic DNA (gDNA) was extracted from 10 mL of peripheral blood by the method of Sambrook et al. (16).

Northern analysis of poly(A)+ mRNA. Poly(A)+ mRNA was electrophoresed in a 1.0% denaturing agarose gel (agarose from Progen, Darra, Australia), transferred to a nylon membrane (Genescreen Plus; Dupont, Boston, MA), and baked for 2 h in a reduced-pressure oven (Scientific Equipment Manufacturers, South Australia). The Northern transfer was hybridized with a 32P-labeled full-length UROD gene cDNA probe (a gift from H. de Verneuil of Bordeaux University, France), which was labeled by the random hexamer priming method (with random hexamers from Pharmacia) (16). Hybridization was performed at 60°C in 1 mmol/L sodium chloride, 1 g/L sodium dodecyl sulfate (SDS), and 100 g/L dextran sulfate (Pharmacia). After hybridization, the membrane was washed once in 2× standard saline citrate (SSC) for 5 min, and two times (15 min each) in 0.1× SSC containing 1 g/L SDS. All washings were performed at 60°C.

Restriction endonuclease digestion on genomic DNA. gDNA was digested with BamHI and EcoRI (Promega, Madison, WI) and subjected to electrophoresis in a 1.0% agarose gel, followed by Southern blotting with a nylon membrane (Genescreen Plus). The nylon membrane was hybridized with a 32P-labeled full-length UROD gene cDNA probe as described above, except that the hybridization and washings were performed at 65°C.

PCR amplification of DNA. cDNA was amplified by two rounds of the PCR (17). Between 5% and 20% of the synthesized cDNA was used in separate reactions for the first round of PCR amplification; 10% of the generated first-round PCR products was used in the second round of PCR amplification. The 5′ and 3′ oligonucleotide primers for the first round were BM2 and BM1, respectively (Table 1). The internal oligonucleotide primers for the second round of PCR amplifications, N1 and N4, were synthesized with flanking EcoRI and HindIII sites, respectively, which were used for subsequent cloning (Table 1). PCR amplifications were performed in an Innovonics Genemachine (Bartelt, Melbourne, Australia) or a Hybrid Thermal Reactor (Hybaid, Teddington, UK) with a thermostable polymerase enzyme (Taq polymerase; Promega), 30 cycles per round (round 1: 2 min at 94°C, 2 min at 55°C, and 2 min at 72°C; round 2: 2 min at 94°C, 2 min at 50°C, and 2 min at 72°C). The choice of 2 min at 72°C for the primer extension was based on the expected size of the PCR product, 1133 bp, and experimental evidence that 1 U of Taq polymerase

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence</th>
<th>Site (nt position)</th>
<th>Ts (°C)</th>
<th>Predicted size, bp</th>
<th>Product</th>
</tr>
</thead>
<tbody>
<tr>
<td>BM2</td>
<td>5'-gcagccaaaagacagcgtggctgcc-3'</td>
<td>5'-UTR (1-18)</td>
<td>55</td>
<td>1173</td>
<td>Full cDNA</td>
</tr>
<tr>
<td>BM1</td>
<td>5'-gcagccaaaagacagcgtggctgcc-3'</td>
<td>3'-UTR (1154-1173)</td>
<td>55</td>
<td>1133</td>
<td>Full cDNA</td>
</tr>
<tr>
<td>N1</td>
<td>5'-gcagccaaaagacagcgtggctgcc-3'</td>
<td>Exon 1 (19-38)</td>
<td>50</td>
<td>400</td>
<td>5'-int 1</td>
</tr>
<tr>
<td>N4</td>
<td>5'-gcagccaaaagacagcgtggctgcc-3'</td>
<td>Exon 2 (105-125)</td>
<td>55</td>
<td>450</td>
<td>3'-int 1</td>
</tr>
<tr>
<td>BM2</td>
<td>5'-gcagccaaaagacagcgtggctgcc-3'</td>
<td>Exon 2 (29-59)</td>
<td>55</td>
<td>340</td>
<td>Int 2</td>
</tr>
<tr>
<td>BM2</td>
<td>5'-gcagccaaaagacagcgtggctgcc-3'</td>
<td>Exon 3 (208-227)</td>
<td>55</td>
<td>500</td>
<td>Int 3 &amp; 4</td>
</tr>
<tr>
<td>BM2</td>
<td>5'-gcagccaaaagacagcgtggctgcc-3'</td>
<td>Exon 4 (359-354)</td>
<td>55</td>
<td>340</td>
<td>Int 5</td>
</tr>
<tr>
<td>N2</td>
<td>5'-gcagccaaaagacagcgtggctgcc-3'</td>
<td>Exon 6 (535-572)</td>
<td>55</td>
<td>700</td>
<td>Int 6</td>
</tr>
<tr>
<td>N3</td>
<td>5'-gcagccaaaagacagcgtggctgcc-3'</td>
<td>Exon 5 (492-511)</td>
<td>55</td>
<td>400</td>
<td>Int 7 &amp; 8</td>
</tr>
<tr>
<td>N4</td>
<td>5'-gcagccaaaagacagcgtggctgcc-3'</td>
<td>Exon 9 (913-930)</td>
<td>55</td>
<td>500</td>
<td>Int 9</td>
</tr>
<tr>
<td>N4</td>
<td>5'-gcagccaaaagacagcgtggctgcc-3'</td>
<td>Exon 9 (854-952)</td>
<td>55</td>
<td>400</td>
<td>Int 9</td>
</tr>
<tr>
<td>E10</td>
<td>5'-gcagccaaaagacagcgtggctgcc-3'</td>
<td>Exon 10 (988-1007)</td>
<td>55</td>
<td>400</td>
<td>Int 9</td>
</tr>
</tbody>
</table>

UTR, untranslated region; int, intron; nt, nucleotide; nt position 376 is site of base change (ME4: see text).

* First entry of each pair is the sense orientation; second entry is the antisense.

** Ts, annealing temperature (°C).

*** From ref. 12.
can incorporate 10 nmol of dNTPs in 30 min at 74°C and therefore extend at a rate of 1000 bases/min (18).

gDNA was amplified by PCR with the primer pairs shown in Table 1. For all primer pairs, PCR was conducted for 40 cycles, each consisting of 2 min at 95°C, 2 min at 55°C, and 2 min at 72°C. Then 10% of the PCR products was electrophoresed in 1.5% agarose, transferred by the method of Southern onto a nylon membrane, and hybridized with a 32P-labeled full-length cDNA probe as described above.

A negative control, consisting of the entire PCR reaction mixture but without gDNA or cDNA, was included in every PCR experiment and was confirmed as negative as assessed by hybridization with the radiolabeled UROD cDNA probe.

Cloning and sequencing PCR products. Positively hybridizing PCR products were either extracted with phenol/chloroform (16) and precipitated or electrophoresed in low-melting-point agarose (USB Corp., Cleveland, OH). The appropriate band or bands were excised from the agarose gel, and the PCR products were extracted by Magic Prep (Promega) according to the manufacturer's instructions. The amplified DNA was then sequenced directly or digested (with EcoRI and HindIII) and ligated into pUC13 (Promega) or blunt-ended and ligated into pBluescript (Stratagene, La Jolla, CA). At least six independent clones were characterized for indirect sequencing of PCR products. Direct and indirect sequencing of alternative transcripts was confirmed in both directions by using a modified version of the dideoxy method (19) with either Sequenase (USB Corp.) or T7 DNA polymerase (Pharmacia).

Screening for UROD mRNA transcripts lacking exon 6. After synthesizing an 18-bp oligonucleotide probe that spanned the exon 5/exon 7 junction (JM5: GGTGCGC CCCAGCATGGCAG), we end-labeled the oligonucleotide with γ-32P-ATP (Bresatec), removing the free 32P-ATP with a NAP-5 column containing Sephadex G-25 medium (Pharmacia). Products of the PCR amplification reactions were hybridized as described above, at 60°C. The membrane was washed three times in 5× SSC containing 1 g/L SDS at room temperature for 15 min, then once for 15 min at 60°C in 2× SSC/1 g/L SDS.

Results
Structure and Function of the UROD Gene and Its Product

Digestion of gDNA with BamHI and EcoRI followed by Southern analysis confirmed a normal configuration of the UROD gene in the normal individuals and in the fPCT patients with no evidence of deletions or rearrangements (Fig. 1). This was supported by the data from Northern analysis and the assays of protein function. Analysis of poly(A) mRNA from leukocytes revealed only the predicted 1.2-kb UROD mRNA and precursor mRNA in normal subject #1 and the fPCT patients. The failure of Northern analysis to demonstrate the presence of any smaller UROD mRNA transcripts suggested that they were not abundant (Fig. 2).

To study UROD function, we established a reference range for erythrocyte UROD activity from results for 20 healthy laboratory workers: 1.8–4.0 U/L. The erythrocyte UROD activity in the two normal individuals analyzed in this study was within the reference range 3.1 and 2.4 U/L, respectively, whereas that of the fPCT patients was ~50% of normal (1.2 U/L for the proband and 1.3 U/L for her daughter).
Multiple UROD Transcripts

Electrophoresis of PCR-amplified cDNA in agarose gel revealed multiple bands of different sizes, all of which hybridized with a $^{32}$P-labeled full-length cDNA probe (Fig. 3). The predicted size of the full-length PCR product was 1133 bp; however, the PCR products varied from 350 to 1133 bp and were observed both in the normal individuals and in the patients with PCT. Cloning and subsequent sequencing of these products revealed several alternative transcripts for UROD (Fig. 4). All transcripts, including the full-length UROD transcript, were observed only after two rounds of PCR. The full-length UROD mRNA (transcript A) was observed in both fPCT patients and normal #1 but not in normal #2. This can be explained by the nature of PCR; namely, that small sequences of DNA are preferentially amplified over large ones. This phenomenon was probably enhanced in the study by our performing two rounds of PCR. Transcript B, which lacked exon 6, was identified by sequence analysis in the proband. Transcript B was also detected in both normal subjects by hybridization of the Southern transfer of the PCR products with the oligonucleotide JMS (data not shown). Subsequent nucleotide sequence analysis confirmed an exon 5 to 7 splicing event in normal #2. Transcript C was identified in only one normal individual; transcripts D and E (lacking four and five exons, respectively) were detected only in the proband. Transcript F was observed in the proband's daughter, whereas both transcripts F and G (each lacking seven exons) were observed in the proband and normal #1. Transcript G also retained 50 bp of the 3' end of intron 1. In this transcript, a cryptic splice site 5'-CGCGCTTTCCCCAG-3' was utilized instead of the usual splice site 5'-CACCT-GATGCCCAG-3'.

Transcripts C, D, and F resulted in a shift in the reading frame and the introduction of a stop codon; these, therefore, would not yield a functional UROD protein. Transcript G had a stop codon midway through the retained portion of intron 1 and therefore would produce a predicted protein of only 12 amino acids. Transcripts B and E potentially encoded proteins that lacked amino acids 159–212 and 93–314, respectively, but retained the amino and carboxyl terminals of the UROD protein. Given that neither the mode of action of the UROD enzyme nor the location of its catalytic sites is known, the impact of these truncated proteins on UROD function is unknown. Except for transcript G, which utilized a cryptic splice site in intron 1, all PCR products in which the nucleotide sequence was determined contained exons that abutted the previously defined exon boundaries.

Characterization of Intron/Exon Boundaries of UROD Gene

All intron/exon boundaries of the UROD gene were characterized and found to be identical in the two normal subjects and the proband. The splice donor and acceptor sequences complied with the known consensus sequences that flank intron/exon boundaries (5) and were identical to those previously published for the
UROD gene with two exceptions: a single-base substitution in the exon 10 acceptor site and a single-base deletion in the exon 3 acceptor site. At the 3' end of intron 9 (the exon 10 acceptor site), a T→C substitution was observed, TGTTGTCCCCTAG/GAG to TGTTG-TGTCCCCAG/GAG. The deletion of an adenine at the 3' end of intron 2 (the acceptor site of exon 3) was also observed in all individuals: CCTCTGTATGCAG/AGT. These two changes were also observed in introns 2 and 9 of an independently derived UROD genomic clone that we have sequenced (data not shown). As discussed later, these changes are unlikely to be responsible for the multiple alternative transcripts we observed. In addition, characterization of the intron/exon junctions failed to identify any error at the genomic level to explain the mRNA lacking exon 6.

Single-Base Differences

Four additional deviations from the initial published sequence of the UROD gene have been described (7, 8) and all were observed in this study. They occur at positions 87, 325, 931, and 1027 in the coding region and include a T→C change at positions 87 (GCT→GCC), 931 (TG→CTG), and 1027 (TTG→CTG)—none of which changes the encoded amino acids, alanine or leucine. However, at position 325, an A→G substitution (AGC→GCG) changes the amino acid deduced from the initial cDNA sequence from serine to glycine. This glycine residue has been demonstrated by direct protein sequencing (5); it was also confirmed by our nucleotide sequence analysis, which showed this A→G change in all individuals (6, 8). A deviation not previously described was observed at positions 376, 377, and 378: GCG→CGC. All of these deviations from the initial published sequence were observed in all clones analyzed in this study.

Discussion

We describe here alternative splicing of mRNAs in the UROD gene that result in transcripts lacking between one and seven exons. Such splicing was observed both in normal individuals and two patients with IFCT. Moreover, analysis of the UROD gene splice site sequences in the two normal subjects and the proband failed to explain these alternative transcripts.

Studies of the UROD gene have not greatly improved our understanding of the cause of IFCT at the molecular level. A single-point mutation in the coding region has been described at position 860 that results in the substitution of glycine for valine (7). In addition, an mRNA lacking exon 6 has been observed, attributed to a G→C substitution at the junction of exon 6 and intron 6: CAG/GT GAGT to CAG/GT GAGT. This abnormality was observed by Garey et al. (12) in about half of the PCR-generated clones analyzed after a single round of PCR amplification. The same abnormality was found in 5 of 22 unrelated families with IFCT when an oligonucleotide probe containing the G→C base substitution was used (12).

Several lines of evidence suggest that the alternative UROD gene transcripts described here were not artificial. First, in six of the seven transcripts, the usual splice sites were utilized, joining the UROD exons at the previously defined exon boundaries; the exception involved a cryptic splice site in intron 1. Second, not all of the transcripts defined were observed in all individuals, which makes artifacts due to PCR contamination unlikely. This further suggests that additional alternative forms of the UROD mRNA may exist. The transcripts we found may represent a small percentage failure rate of the normal splicing mechanism or perhaps a posttranscriptional modification. In any event, the abundance of these transcripts appears to be low. The likely explanation for their detection is the preferential selection of small DNA fragments for amplification by PCR, which is further enhanced by performing two rounds. Perhaps smaller transcripts could be detected in other genes that might represent the byproducts of normal splicing.

The impact of these transcripts on UROD protein function, if any, is unknown. Transcripts C, D, and F are unlikely to encode a functional UROD protein because there is no long open reading frame. Transcript G potentially could encode a peptide truncated at the amino terminus through a stop codon introduced by the retention of part of intron 1. Transcripts B and E both had an open reading frame and could potentially yield proteins that lacked amino acids 159–212 and 93–314, respectively. Garey et al. (12) also described transcript B; through functional studies they demonstrated that the encoded protein lacked catalytic activity and that the mRNA was rapidly degraded.

Although we observed two deviations from the published UROD gene splice sequences, i.e., in the acceptor sites of exons 3 and 10, they are unlikely to have had any bearing on the alternative splicing events. The AG dinucleotide, which is conserved in >99% of intron/exon junctions studied (20), was present at both sites. At the −3 position (the site of the T→C base alteration in intron 9), C occurs in 65% of splice junctions and T in 31%. At position −10, adenine is seen in only 15% of cases, and its deletion in intron 2 resulted in a thymine residue at this position, as observed in 49% of splice sites. From the results of these splice junction studies, the deviations observed in our subjects would probably have no effect on normal splicing.

Precedent exists for alternative splicing of other genes. Alternative transcripts have been associated with point mutations in the splice site consensus sequences, especially those that flout the GT→AG rule (21–23). In the analysis of 139 splice junctions of nuclear and viral genes, A occupied the first position of every intron (the donor site of every exon) (20). However, several normally splicing genes possess splice sequences that fail to comply with the GT→AG rule. In most of these nonconsensus splice sites GT is replaced by GC. In such instances, the nucleotides that surround the GC dinucleotide in the splice junction show greater homology to the prototype consensus splice sequence and thus to U1 RNA than the average splice sequence (24), and probably allow normal splicing to proceed. In
vitro studies have demonstrated that this substitution simply reduces the splicing efficiency (25). Alternative splicing has also been observed in some genes without any errors in the appropriate splice site junctions. For example, in the respiratory epithelium of normal individuals, an mRNA lacking exon 9 was observed in the cystic fibrosis transmembrane conductance regulator gene (26). In addition, alternative mRNA splicing of the platelet-derived growth factor A chain gene has been demonstrated, where either exons 3 or 6, or both, were deleted (27). In both examples, PCR was used.

The finding of these aberrant transcripts in normal individuals indicates that their presence is clinically and biochemically insignificant. The identification of transcript B in both normal individuals and the proband, and without any mutations in the splice junctions of the gene, suggests that the UROD mRNA lacking exon 6 is unlikely to be responsible for the PCT in this patient. In addition, investigators should be cautious about drawing conclusions regarding the cause of PCT from PCR analysis of UROD transcripts. It is probably necessary to characterize the appropriate splice site in the gDNA. These findings may have implications not only for interpretation of alternative UROD transcripts but also for transcripts observed in other genes studied by PCR.

We thank Hubert de Verneuil of Bordeaux University, France, for the gift of the cloned UROD cDNA. We acknowledge George Varigos of the Royal Melbourne Hospital Dermatology Department for allowing us to investigate his patients and Dennis Blake for performing the porphyrin analyses and for rigorous follow-up of our porphyria patients. This work was in part supported by the Royal Melbourne Hospital Victor Hurley Fund and the National Health and Medical Research Council, Canberra, Australia.

References

25. Aebi M, Hornig H, Weissmann C. 5' Cleaveage site in eukaryotic pre-mRNA splicing is determined by the overall 5' splice region, not by the conserved 5' GU. Cell 1987;50:237–46.