Uroporphyrinogen III Synthase

AN ALTERNATIVE PROMOTER CONTROLS ERYTHROID-SPECIFIC EXPRESSION IN THE MURINE GENE*

(Received for publication, April 2, 1999, and in revised form, November 8, 1999)

Gerardo I. Aizencang, David F. Bishop, Douglas Forrest, Kenneth H. Astrin, and Robert J. Desnick‡

From the Department of Human Genetics, Mount Sinai School of Medicine, New York, New York 10029

Uroporphyrinogen III synthase (URO-synthase, EC 4.2.1.75) is the fourth enzyme of the heme biosynthetic pathway and is the defective enzyme in congenital erythropoietic porphyria. To investigate the erythroid-specific expression of murine URO-synthase, the cDNA and –24-kilobase genomic sequences were isolated and characterized. The alternative transcripts were identified containing different 5′-untranslated regions (5′-UTRs), but identical coding exons 2B through 10. Transcripts with 5′-UTR exon 1A alone or fused to exon 1B were ubiquitously expressed (housekeeping), whereas transcripts with 5′-UTR exon 2A were only present in erythroid cells (erythroid-specific). Analysis of the TATA-less housekeeping promoter upstream of exon 1A revealed binding sites for ubiquitously expressed transcription factors Sp1, NF1, AP1, Oct1, and NRF2. The TATA-less erythroid-specific promoter upstream of exon 2A had nine putative GATA1 erythroid enhancer binding sites. Luciferase promoter/reporter constructs transfected into NIH 3T3 and mouse erythroblast leukemia cells indicated that the housekeeping promoter was active in both cell lines, while the erythroid promoter was active only in erythroid cells. Site-specific mutations of the first GATA1 binding site markedly reduced luciferase activity in K562 cells (<5% of wild type). Thus, housekeeping and erythroid-specific transcripts are expressed from alternative promoters of a single mouse URO-synthase gene.

The genomic sequences encoding the eight mammalian (human and/or murine) heme biosynthetic enzymes have been isolated and characterized (1–7), with the notable exception of URO-synthase, the fourth enzyme in the pathway (URO-synthase; hydroxymethylbilane hydrolyase (cyclizing), EC 4.2.1.75). The genes encoding the first three enzymes have erythroid-specific transcripts generated by a separate gene (5-aminolevulinate synthase) (8) or by alternative promoters of a single gene (5-aminolevulinate dehydratase and hydroxymethylbilane synthase) (2, 9). Their erythroid-specific promoters and enhancers all have GATA1, NF-E2, and CACCC elements that bind transcriptional factors to facilitate the high level expression of these transcripts in erythroid cells. Of the remaining five genes in the pathway, each has been shown to be a single gene with a single promoter containing both erythroid-specific binding elements for enhanced erythroid expression and housekeeping elements for constitutive expression in all cells (6, 10–12), with the notable exception of the URO-synthase gene, which has not been isolated or characterized; nor has its erythroid-specific expression been investigated.

URO-synthase is responsible for the conversion of the linear tetrapyrrole, hydroxymethylbilane, to uroporphyrinogen III, the first cyclic tetrapyrrole and physiologic precursor of heme (13). The enzyme functions as both an isomerase and a cyclase as it catalyzes the intramolecular rearrangement of the pyrrole ring D and tetrapyrrole ring closure of the linear tetrapyrrole, respectively (14, 15). In the absence of URO-synthase activity, hydroxymethylbilane is nonenzymatically cyclized to form the nonphysiologic uroporphyrinogen I isomer, which is then oxidized to uroporphyrin I, a nonmetabolizable and pathogenic compound. URO-synthase has been purified to homogeneity from human erythrocytes, rat liver, and bacteria, yeast, and rat for a review, see Ref. 23, and their significant homologies provided further evidence for a single URO-synthase gene in prokaryotes as well as in the animal kingdom.

The markedly deficient, but not absent, activity of URO-synthase is the enzymatic defect in congenital erythropoietic porphyria (CEP), an inborn error of heme biosynthesis that occurs in humans and cattle (24–26). The enzymatic defect results in the erythroid accumulation of uroporphyrinogen I, which leads to the clinical manifestations. In humans, the

*This work was supported in part by grants from the National Institutes of Health, including Grants 5 R01 DK26824 and 2 R01 DK40885, Grant 5 M01 RR0071 for the Mount Sinai General Clinical Research Center, and Grant 5 P50 HD28822 for the Mount Sinai Child Health Research Center. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

†To whom correspondence should be addressed: Dept. of Human Genetics, Mount Sinai School of Medicine, Fifth Ave. at 100th St., New York, New York 10029. Tel.: 212-659-6700; Fax: 212-360-1809; E-mail: rjdesnick@msvax.mssm.edu.

‡The abbreviations used are: URO-synthase, uroporphyrinogen III synthase; AP2, CLONTECH Marathon® nested adaptor primer 2; CEP, congenital erythropoietic porphyria; EST, expressed sequence tag; Inr, transcription initiation region; MEL, mouse erythroblast leukemia cell line; PCR, polymerase chain reaction; PAGE, polyacrylamide gel electrophoresis; P1, bacteriophage P1; RACE, rapid amplification of cDNA ends; RT-PCR, reverse transcriptase-PCR; UTR, untranslated region; bp, base pair(s); kb, kilobase pair(s); nt, nucleotide(s).
Erythroid-specific Expression of Mouse URO-synthase

Isolation of isolated tissue mRNAs. The RT-PCR products (10 cycles of amplification of exon 1A (primer 2), 1B (primer 4), and antisense primer 17 (Table I; cDNA bp 503–529, numbered according to the human and murine URO-synthase genomic sequence and to determine if mouse URO-synthase had erythroid-specific regulation. In this paper, we report that the mouse URO-synthase gene has distinct erythroid and housekeeping promoters that generate three different transcripts, one being erythroid-specific. In addition, the genomic organization, promoter sequences, and 5′ transcription start sites of the alternative transcripts were determined, and erythroid-specific expression was demonstrated in MEL and human K562 cells.

**EXPERIMENTAL PROCEDURES**

**Rapid Amplification of 5′ cDNA Ends (5′-RACE)—** 5′-RACE was performed using Marathon-Ready® mouse spleen cDNA (CLONTECH, Palo Alto, CA) according to the manufacturer’s procedure. For the PCR, the use of the Marathon Marathon-ready adaptor primer (AP2; CLONTECH, Palo Alto, CA) and the mouse URO-synthase antisense primer 17 (Table I; cDNA bp 503–529, numbered according to the full-length housekeeping URO-synthase cDNA sequence; GenBank™ accession no. U18867.3). The reaction conditions were as follows: 95 °C for 12 min followed by 35 cycles of amplification (94 °C for 30 s and 70 °C for 2 min) and then a final extension step at 70 °C for 7 min. The PCR products were then subcloned into the pGEM-T Vector (Promega, Madison, WI). Positive clones were sequenced with a fluorescent automated DNA sequencer (ABI Prism™ model 377- XL; Perkin-Elmer Applied Biosystems, Inc, Foster City, CA.) using dRhodamine dye terminator chemistry.

**RT-PCR Analysis of Mouse URO-synthase mRNAs—** Total RNAs from different mouse tissues were isolated using the “Ultraspec DNA Isolation System” (Biotecx Laboratories, Inc, Houston, TX). RT-PCR was performed in a final volume of 50 μl using the Titan™ One Tube RT-PCR System (Roche Molecular Biochemicals) according to the manufacturer’s protocol, and the amplification was carried out in a PCR-150 MiniCycler (MJ Research, Inc., Watertown, MA). RT-PCR was initiated at 50 °C for 30 min and then 94 °C for 2 min, followed by 10 cycles of denaturation at 94 °C for 30 s, annealing at 65 °C for 30 s, and extension at 68 °C for 1 min and then 30 cycles of denaturation at 94 °C for 30 s, annealing at 65 °C for 30 s, and extension at 68 °C for 1 min (with each elongation lengthened by an additional 5 s/cycle). The sequences of the sense primers for amplification of exon 1A (primer 2), 1B (primer 4), and 2A (primer 7) are indicated in Table I. Each primer set was designed to produce a PCR product that included at least one intron-exon boundary, thereby eliminating the possibility that DNA contamination was responsible for the resulting products amplified from the isolated tissue mRNAs. The RT-PCR products (10 μl) were subjected to electrophoresis in 1.5% agarose gels and stained with RbBr. The RT-PCR products were purified using the QIAquick PCR Purification Kit (Qiagen, Valencia, CA) and then sequenced using the antisense primer 13 (Table I) to ensure that they corresponded to the targeted URO-synthase genomic sequence.

**Isolation and Characterization of the Mouse URO-synthase Gene—** A bacteriophage P1 DNA library of genomic DNA isolated from mouse strain 129OLA ES cells (Genome Systems, St. Louis, MO) was screened with oligonucleotide primers 4 and 5 (Table I), which amplified an 88-bp fragment of exon 1A/1B and, with primers 10 and 13, which amplified ~600 bp between exons 2 and 3. The conditions for PCR were as follows: 30 cycles of denaturation at 94 °C for 30 s, annealing at 58 °C for 30 s, and extension at 68 °C for 1 min. Three positive clones containing exon 1 and/or exons 2 and 3 were identified and purified. To determine the fragment size, the clones were amplified and sequenced using primers corresponding to the mouse URO-synthase genomic sequence as indicated in Table I. The intron sizes were estimated from the PCR products.

**Sequencing and Analysis of the Housekeeping and Erythroid-specific Promoters—** The mouse URO-synthase genomic P1 clone (mgURO5-1) was directly sequenced using dRhodamine dye terminator chemistry in an automated fluorescent DNA sequencer as described above. To characterize the housekeeping and erythroid promoter regions, the sequenced upstream of exons 1A and 2A were initially sequenced using the respective exonic primers (5 and 9, Table I), and then subsequent sequencing reactions were carried out using oligonucleotide primers based on the most clearly readable 5′ sequence obtained. All sequences were confirmed in both orientations.

The URO-synthase housekeeping and erythroid-specific 5′-flanking promoter/enhancer regions were analyzed using the MatInspector program (32) with the TRANSFAC 3.4 data base of vertebrate transcription factor consensus matrices (available on the World Wide Web) with core similarity of 1.0 and matrix similarity of 0.85.

**Construction of the Promoter Reporter Plasmids—** To construct the housekeeping and erythroid promoter/reporter plasmids, the regions containing the sequence of exons 2B and 3 were amplified and subcloned into the pGL3-Firefly Luciferase Vector (Promega). For amplification of the housekeeping promoter, sense primer 1 (corresponding to mouse genomic sequence nt: 1772 to 1753; Fig. 5) and antisense primer 3 (corresponding to mouse genomic sequence nt: 24–43; Fig. 5) were employed. To amplify the erythroid promoter, sense primer 6 (corresponding to mouse genomic nt: 1524 to 1516; Fig. 6) and antisense primer 8 (corresponding to mouse genomic nt: 59–77; Fig. 6) were used. For cloning purposes, 5′ tails containing an SpeI or SalI restriction site were included in the sense and antisense primers, respectively. After amplification, the PCR products were double digested with SalI and SpeI and cloned into the XhoI–NotI sites of the pGL3-Luciferase Vector (Promega). The plasmids were purified using the Qiagen Plasmid Maxi Kit (Qiagen, Valencia, CA) and sequenced to confirm their authenticity. The housekeeping and erythroid promoter/reporter constructs were designed mHP and mEP, respectively.

To test the functionality of the putative erythroid-specific promoter region, a point mutation at T−65 was introduced into the first putative GATA1 site (5′-TTATCA-3′→5′-TTACCA-3′) in the erythroid promoter/reporter construct mEP using the QuickChange™ site-directed mutagenesis kit (Stratagene, La Jolla, CA), according to the manufacturer’s instructions. Amplification was performed with sense primer 30 and antisense primer 31 (both corresponding to mouse erythroid promoter coordinates nt: 79 to −51; Fig. 6). The mutant construct was designated mEP−65C. Sequence analysis confirmed the authenticity of the mutant construct.

**Cell Culture, Transfections, and Luciferase Assays—** Mouse MEL cells (a gift from Dr. George Atweh, Mount Sinai School of Medicine, New York) and human erythroleukemia K562 cells were grown in RPMI 1640 medium. Mouse fibroblast NIH 3T3 cells were grown in Dulbecco’s modified Eagle’s medium. Both media were supplemented with 10% fetal bovine serum, 2 mM l-glutamine, 100 units/ml penicillin, and 100 mg/ml streptomycin (Life Technologies, Inc., Gaithersburg, MD.). All transfections were performed in 12-well plates. For NIH 3T3 cell transfections, 1 × 10⁵ cells were plated 10–16 h prior to transfection. The respective reporter construct (2 μg) and 50 ng of the internal control, Renilla Luciferase-TK (pRL-TK) vector were transfected using 3.2 μl of Lipofectamine and 5 μl of Plus Reagent (Life Technologies). After transfection (24 h), the cells were washed twice with phosphate-buffered saline, harvested, and lysed with 100 μl of Passive Lysis Buffer (Promega Corp.). For MEL and K562 cells, 8 × 10⁵ cells were transfected as above with 2 μg of the reporter construct and 50 ng of the pRL-TK vector, using 4 μl of DMRIE-C reagent (Life Technologies). After transfection (48 h), the cells were collected by centrifugation, washed with phosphate-buffered saline, and lysed with 100 μl of Passive Lysis Buffer. When the MEL cells were induced to differentiate, 2% M hemin was added to the culture medium 24 h prior to transfection. The Renilla and firefly luciferase activities were determined in 20 μl of the cell lysates using the Dual-Luciferase™ Reporter Assay System (Promega) according to the manufacturer’s procedure, with a Microtiter® Plate Luminoimeter (DYNEX Technologies, Inc., Chantilly, VA). The firefly luciferase activities were normalized for...
transfection efficiency using the Renilla luciferase activity as an internal control, and the data were expressed as relative light units. A negative control using the pGL-3 basic vector was included in all experiments. Data represent the means ± S.D. of at least five independent transfection experiments.

RESULTS

5′-RACE of Mouse Splenic cDNA Demonstrates Multiple URO-synthase 5′-UTRs—To assess the possible existence of separate housekeeping and erythroid-specific transcripts for the mouse URO-synthase, 5′-RACE was performed using Marathon-Ready™ mouse spleen cDNA (for details, see "Experimental Procedures"). A, schematic representation of the amplification strategy. The white box represents the URO-synthase cDNA, with the ligated adaptor regions indicated as black boxes. The arrows indicate the approximate positions of the primers used (AP2 and 17). B, of the over 200 RACE clones, seven group 1A and 17 group 2A clones were sequenced. The lengths of the unique 5′-UTRs—

The Mouse URO-synthase Gene Has 10 Exons and Three Alternative 5′-UTRs—Screening of a mouse genomic P1 library (Fig. 1B), schematic representation of the amplification strategy. The white box represents the URO-synthase cDNA, with the ligated adaptor regions indicated as black boxes. The arrows indicate the approximate positions of the primers used (AP2 and 17). B, of the over 200 RACE clones, seven group 1A and 17 group 2A clones were sequenced. The lengths of the unique 5′-UTRs—

\[ \text{transfection efficiency using the Renilla luciferase activity as an internal control, and the data were expressed as relative light units. A negative control using the pGL-3 basic vector was included in all experiments. Data represent the means ± S.D. of at least five independent transfection experiments.} \]

\[ \text{RESULTS} \]

\[ \text{5′-RACE of Mouse Splenic cDNA Demonstrates Multiple URO-synthase 5′-UTRs—To assess the possible existence of separate housekeeping and erythroid-specific transcripts for the mouse URO-synthase, 5′-RACE was performed using Marathon-Ready™ mouse spleen cDNA (for details, see "Experimental Procedures"). A, schematic representation of the amplification strategy. The white box represents the URO-synthase cDNA, with the ligated adaptor regions indicated as black boxes. The arrows indicate the approximate positions of the primers used (AP2 and 17). B, of the over 200 RACE clones, seven group 1A and 17 group 2A clones were sequenced. The lengths of the unique 5′-UTRs—}\]

\[ \text{The Mouse URO-synthase Gene Has 10 Exons and Three Alternative 5′-UTRs—Screening of a mouse genomic P1 library produced three clones positive for URO-synthase primers amplifying both the 5′-UTR 1B region and coding exons 2 and 3. Restriction analysis of the three P1 clones revealed similar band patterns on agarose gel electrophoresis, indicating that the three clones contained common regions of genomic DNA. Therefore, the genomic organization of the clone with the largest insert, mgUROS-1, was determined. Amplification and sequencing of the mgUROS-1 P1 DNA using primers designed from the 5′-UTR sequences 1A, 1B, and 2A and the first region of common sequence (designated 2B) in the 1A, 1B, and 2A-containing transcripts facilitated determination of the 5′ region of the gene (Fig. 2A). As shown in Fig. 2, the 5′-UTR sequence previously reported for the mouse liver URO-synthase cDNA (22) was 19 bp downstream from the group 1A 5′-UTR sequences. The group 2A 5′-UTR sequences were located ~6 kb downstream of the group 1A 5′-UTR gene sequences and were directly upstream of the common exonic region containing the translation start site, which was designated exon 2B. Therefore, the group 1A 5′-UTR sequence was designated exon 1A and the directly adjacent downstream 1B 5′-UTR sequence was designated exon 1B. Analysis of the longest URO-synthase cDNA clone (GenBank™ accession no. U16216) revealed that it contained the entire 85-bp exon 1B, 21 bp of exon 1A, and 41 bp at the 5′-end, which was an artifact derived from the murine mitochondrial 16 S rRNA. As shown in Fig. 2B, the mitochondrial 16 S rRNA sequence shared a 7-bp overlapping region with the 5′-end of the URO-synthase cDNA, which contains a BanHI restriction site. Most likely, this artifactual sequence resulted from concatamerization of the Sau3AI partial digests used to construct the library. Computer-assisted searches of GenBank™ revealed additional examples of mouse URO-synthase cDNAs including alternatively spliced sequences for exons 1A and 1B as well as a fetal mouse EST containing the exon 2A extension of exon 2B (Fig. 2B).} \]

\[ \text{Amplification and sequencing of the exonic regions in the mouse URO-synthase gene using primers based on the cDNA sequence (Table I) demonstrated that the P1 clone contained the entire ~24-kb URO-synthase gene with 10 exons. The exon/intron boundaries were sequenced, and the sizes of the introns were estimated from PCR products (Table II). The introns ranged in length from approximately 0.5 to 7.5 kb, while the exons ranged from 75 to 761 bp, and their sequences matched that of the previously reported cDNA sequence (22). The sequences at all of the intron-exon boundaries conformed to the GT-AG rule (33).} \]

\[ \text{Tissue-specific Expression of the Alternative URO-synthase Transcripts—To study expression of the URO-synthase alternative transcripts in different mouse tissues, RT-PCR was performed with total RNA from whole fetus (embryonic day 10–12) and adult brain, kidney, liver, and spleen using sense primers designed to bind specifically to exon 1A (primer 2), the alternative splice junction region of exon 1A and 1B (primer 4), or exon 2A (primer 7; Fig. 3A). Amplification with exon 1A-specific primer 2 generated two products, which were sequenced (Fig. 3A). The smaller RT-PCR product was 484 bp and contained exon 1A spliced to exon 2B, while the longer product was 569 bp and had exons 1A and 1B (designated exon 1A/1B) fused to exon 2B. RT-PCR products with alternatively spliced exon 1A or exon 1A/1B were detected in all tissues examined, but the relative amount of the exon 1A/1B product was reduced in brain and fetal tissue (Fig. 3C). Primer 4 yielded a 528-bp RT-PCR product, which contained exon 1A/1B and was present in all tissues studied but again was reduced in brain and fetal tissue (Fig. 3B), indicating that the exon 1A/1B-containing transcripts were ubiquitously expressed, presumably from a housekeeping promoter.} \]

\[ \text{In contrast, a markedly different expression pattern was}\]
observed for exon 2A-specific transcripts. A single 660-bp RT-PCR product (Fig. 3D) was detected only in spleen, whole fetus, and G1/ER cell RNA, the latter RNA from a murine erythroid proerythroblast cell line that was induced by GATA1 to undergo terminal erythroid differentiation (34) (a generous gift from Dr. Mitchell Weiss, Children’s Hospital, Harvard Medical School, Boston, MA). Exon 2A was not detected in nonerythroid tissues even after the PCR conditions were modified to detect
Total RNA was extracted from mouse total fetus (F), brain (B), kidney (K), liver (L), spleen (S), and the ES cell line G1/ER (G). M represents size markers, and C represents a control of no template DNA. RT-PCR was performed as detailed under “Experimental Procedures” using sense primers specific for exon 1A, exon 1AB, or exon 2A as indicated in A. The PCR products were separated in a 1.5% agarose gel and stained with ethidium bromide. A, schematic representation of the three alternative URO-synthase transcripts showing the location of the primers and the predicted sizes of the RT-PCR products; B, RT-PCR using primer 4; C, RT-PCR using primer 2; D, RT-PCR using primer 7.

**Fig. 3. RT-PCR of total RNA from different mouse tissues.** Total RNA was extracted from mouse total fetus (F), brain (B), kidney (K), liver (L), spleen (S), and the ES cell line G1/ER (G). M represents size markers, and C represents a control of no template DNA. RT-PCR was performed as detailed under “Experimental Procedures” using sense primers specific for exon 1A, exon 1AB, or exon 2A as indicated in A. The PCR products were separated in a 1.5% agarose gel and stained with ethidium bromide. A, schematic representation of the three alternative URO-synthase transcripts showing the location of the primers and the predicted sizes of the RT-PCR products; B, RT-PCR using primer 4; C, RT-PCR using primer 2; D, RT-PCR using primer 7.

rare transcripts, supporting the finding that the alternative exon 2A-containing transcript was expressed only in erythroid cells. Thus, the URO-synthase gene contains alternatively spliced exons (1A, 1A/1B, and 2A), which encode different 5'-UTRs that are expressed in the housekeeping or erythroid-specific transcripts (Fig. 4).

**Sequence Analysis of the Housekeeping Promoter/Enhancer Region**—To characterize the “housekeeping” promoter/enhancer region, a 1772-bp region immediately upstream of exon 1A/1B was sequenced from the URO-synthase P1 genomic clone (Fig. 5; GenBank™ accession no. AF133257). The proximal promoter region lacked a TATA box; however, there was a multiplicity of transcription start sites as indicated by the multiple 5'-RACE ends and cDNA clones (Figs. 1B, 2B, and 5). According to convention (e.g. see Ref. 35), the cap site (position +1) was designated the most 5’ base in the longest murine URO-synthase cDNA clone (GenBank™ accession no. AA980856), which was only 20 bp longer than the longest 5'-RACE product (Fig. 5). The full-length URO-synthase cDNA containing the exon 1A/1B 5'-UTR is available (GenBank™ accession no. U188672).

Computer-assisted analysis of the putative promoter/enhancer elements in the 1772-bp upstream of exon 1A/1B identified the following putative cis-acting sequences. These transcription factor binding sites were concentrated in the proximal 350 bp immediately upstream of exon 1A/1B and included NRF2 (nt –8 to –17 and –93 to –102), NF1 (nt –21 to –38), Maf (nt –50 to –57), IkappaB (nt –83 to –88, –150 to –162, –212 to –223, and –314 to –325), Ets-1 (nt –165 to –174, –184 to –193, and –236 to –249), Sp1 (nt –253 to –258), AP1 (nt –276 to –286), and Oct1 (nt –291 to –304). The Sp1, NF1, AP1, Oct1, and NRF2 elements are ubiquitously expressed transcription factor binding sites, consistent with the housekeeping promoter function of this region. No erythroid-specific GATA1, CACCC, or NF-E2 binding sites were present, and only three widely separated CACCC sites were found in the 1400 nt upstream from this region.

**Sequence Analysis of the Erythroid-specific Promoter/Enhancer Region**—The transcription initiation site of the erythroid-specific URO-synthase transcript was based on the 5'-end of the longest cloned cDNA, which was consistent with the sites found by 5’-RACE analysis. Among the erythroid-specific clones, the clone with the longest exon 2A (31 bp) was from a mouse fetal liver library (GenBank™ accession no. AA272198). 5’-RACE analysis revealed 10 different exon 2A initiation sites, ranging in size from 5 to 41 bp, with a median length of 16 bp.

Computer-assisted analysis of the 1534-bp region upstream of the exon 2A cap site revealed a series of erythroid-specific transcription factor binding sites, no consensus TATA box, and three Inr initiation sequences (Fig. 6; GenBank™ no. AF133258). The Inr sequence is a powerful promoter core initiation element that is functionally analogous to the TATA box (36, 37). Transcription initiation typically occurs within the YYA(N/T)A(YY consensus sequence for Inr, most often at the first A (38). A consensus Inr motif (TCACTCT) was located at –3 to –4 in the erythroid promoter region, while two additional tandem Inr motifs were in exon 2 (Inr2 (nt 7–13) and Inr3 (nt 13–19)), each with a single mismatch to the consensus sequence. Of interest, the most frequent 5’-RACE ends were within the Inr2 and Inr3 regions. The assignment of the CAP site was consistent with the following findings: 1) the first nucleotide of the longest cloned cDNA; 2) the predicted transcription start site for the observed Inr element was only 1 bp longer than the URO-synthase clone from the mouse fetal liver library (GenBank™ accession no. AA272198); and 3) all but one RACE end were shorter than the assigned CAP site. In the proximal promoter region, perfect matches to the GATA1 consensus elements were at nt –63, –200, and –383. More distally, one NF-E2, one CACCC, and six GATA1 elements were present between –280 and –1535 bp.

**Functional Analysis of the Housekeeping and Erythroid-Specific Promoters**—Reporter gene constructs were designed to evaluate the function of the putative housekeeping and erythroid-specific promoters in cultured nonerythroid NIH 3T3 cells and erythroid MEL cells. The housekeeping promoter construct, mHPr, had 1772 bp of the genomic DNA upstream of exon 1A/1B as well as the first 21 bp of exon 1A/1B (Fig. 5), fused in front of the firefly luciferase cDNA in the pGL3 vector. The erythroid-specific promoter, mEPr, construct contained 1534 bp upstream of exon 2A as well as the entire exon 2A sequence and the first 46 bp of exon 2B (Fig. 6; all noncoding) in front of the firefly luciferase cDNA in the pGL3 vector.

NIH 3T3 and MEL cells were co-transfected individually with each of the luciferase expression constructs and with an internal control pRL-TK vector containing the HSV-TK promoter upstream of the Renilla luciferase cDNA. Twenty-four h after transfection, the housekeeping promoter was active in both cell lines, while the erythroid promoter was detected only in uninduced and induced MEL cells, as shown in Fig. 7. In NIH 3T3 cells, the activity of the housekeeping promoter was over 80 times higher than that detected for the erythroid promoter. It should be noted that the luminescence measured for the erythroid construct in the NIH 3T3 cells was only 2 times higher than the background level of the pGL3 vector alone (data not shown). In contrast, in uninduced MEL cells, the level of housekeeping promoter expression was similar to that of the erythroid promoter. When the cells were induced with Me2SO to undergo erythroid differentiation, both activities increased;
however, the erythroid promoter activity increased over 4 times higher than that of the uninduced cells, while the housekeeping promoter activity was increased only 1.7 times. It is important to note that a markedly lower transfection efficiency was achieved for MEL cells (0.7%) compared with that (13%) for the NIH 3T3 cells as determined by transfections with a green fluorescence protein vector (data not shown). This fact partially explains the lower luminescence values obtained when MEL cells were transfected with the promoter constructs.

To demonstrate the functionality of the erythroid promoter, a new construct, mEPr-65C, was prepared by introducing a T<sup>265C</sup> point mutation in the first putative GATA1 binding site. This construct was then transfected into MEL cells, and the resulting promoter activity was measured.

FIG. 4. Structure of the murine URO-synthase gene. The intron-exon structure schematic is drawn to scale, with the spliced transcripts enlarged 10-fold. ATG and TGA mark the translation start and stop sites, respectively. The arrows indicate the respective housekeeping and erythroid promoter regions. The three alternative transcripts (1A/1B, 1A, and 2A) differing in their 5’-ends are shown above and below the URO-synthase gene.

FIG. 5. The housekeeping promoter region of murine URO-synthase. This sequence corresponds to nucleotides 1–2242 (GenBank accession number AF133258). Exons 1A and 1B are boxed, with the upward arrows indicating the splice junctions. The 5’-ends of isolated RACE clones are indicated by the fine vertical lines with numbers over each indicating the number of clones found for each length. The cap site is indicated by the horizontal arrow and corresponds to the EST sequence GenBank number AA980856. Consensus binding sites for transcription factors are indicated by underlines and abbreviations for the factor name, with the angle brackets indicating that the site was found on the sense (>) or antisense (<) strand.
site in the erythroid promoter/reporter construct. The mEPr and mEPr-65C constructs were expressed in human K562 cells, and their luciferase activities were determined. Compared with the high level of luciferase activity expressed by the mEPr construct, the mutant EPr-65C construct expressed only 4.3 and 2.8% of wild type activity in uninduced and hemin-induced K562 cells, respectively (Table III).

**DISCUSSION**

Heme is the essential pigment of life and is the prosthetic group in hemoglobin, myoglobin, the cytochromes, and many other hemoproteins. In mammals, approximately 85% of heme synthesis occurs in erythroid cells, indicating the importance of the erythroid-specific production. The deficient activity of the individual enzymes in the heme biosynthetic pathway results in a group of disorders known as the porphyrias, which have been classified as hepatic or erythroid based on their clinical manifestations (39). Previous studies in this (2, 3, 40) and other laboratories (9, 41) demonstrated that the first three enzymes of the heme biosynthetic pathway have unique erythroid-specific and housekeeping transcripts encoded by a single gene or by different genes. However, recent evidence suggested that each of the remaining five genes in the pathway had a single transcript under the control of a single promoter that contained the canonical binding sites for erythroid expression including GATA1, the CACCC-binding protein, and, usually, NE-F2 (6, 10–12, 31). Of these genes, only URO-synthase had not been isolated; nor had its erythroid-specific regulation been investigated.

In this paper, we report that the murine URO-synthase gene has distinct housekeeping and erythroid-specific promoters, the latter controlling its erythroid-specific expression. The existence of the mouse URO-synthase erythroid-specific promoter was suggested by 5'-RACE experiments with splenic total RNA, which revealed two 5'-UTRs (1A and 2A) that differed from the previously reported mouse URO-synthase cDNA sequence (GenBank™ accession no. U16216; Ref. 22). A computer-assisted BLAST search of the mouse EST data base with the three 5'-sequences identified an exon 1A-containing URO-synthase sequence from mouse mammary gland (GenBank™ accession no. AA980856) and an exon 2A URO-synthase cDNA from mouse fetal tissue (GenBank™ accession no. AA272198). The presence of an exon 1A transcript from mammary gland and an exon 2A transcript from fetal tissue supported the notion of housekeeping and erythroid-specific transcripts encoded by a single URO-synthase gene with alternative 5'-UTRs and promoters.

Characterization of the mouse URO-synthase gene from a P1 library revealed an approximately 24-kb sequence containing 10 exons. The exon 1A sequence or the entire exon 1 (1A/1B) was ubiquitously expressed in all tissues studied, and one or the other was alternatively spliced directly to the coding exons, beginning with exon 2B. In contrast, exon 2A was expressed only in erythroid tissues with transcription initiated from a 31-bp 5'-extension of exon 2B, such that the entire erythroid message encoded the same polypeptide as the housekeeping transcript. Thus, the expression of the mouse URO-synthase gene, with unique transcripts resulting from housekeeping and erythroid-specific promoters, was similar to the expression of the human and mouse Δ-aminolevulinate dehydratase (2, 42) and hydroxymethylbilane synthase genes (9, 41, 43). Two pro-
moters generate transcripts in all three genes that differ in their 5'-UTRs; their transcripts encode the identical or essentially identical polypeptides, the housekeeping promoters are located upstream of the first exon, and the erythroid-specific promoters are in intron 1, immediately upstream from exon 2 in each gene. Unique to the URO-synthase gene, exon 1 underlies alternative splicing with either exon 1A or the entire exon 1 (1A/1B), as shown in Figs. 2A and 4. The functional basis of the two housekeeping transcripts is not known.

The functional analysis of the erythroid and housekeeping promoter sequences revealed that the erythroid promoter was active only in erythroid cells, indicating its erythroid specificity (Fig. 7), while the housekeeping promoter was active in erythroid (MEL) and nonerythroid (NIH 3T3) cells. Me2SO-induced erythroid differentiation of the MEL cells increased the activity of both promoters (Fig. 7B). Nevertheless, the erythroid promoter was more responsive to erythroid differentiation, as the ratio of erythroid to housekeeping activity increased from 0.6 to 1.6 (Fig. 7B). Moreover, the murine erythroid promoter/reporter construct was highly active in uninduced and induced human K562 cells (Table III). Of particular relevance, site-specific mutagenesis of the first GATA1 binding element in the murine erythroid promoter markedly reduced promoter/reporter activity in both uninduced and induced K562 cells. The fact that the wild type construct was highly expressed in these human erythroleukemia cells and the fact that the GATA1 site was functional in both murine and human erythroid cells.

Transcriptional activities of the housekeeping and erythroid promoters during erythroid differentiation also were similar to each other for the 5-aminolevulinate dehydratase and the hydroxymethylbilane synthase genes (43, 44). Of note, however, the ratio of the URO-synthase erythroid to housekeeping 5'-UTRs in the mouse spleen was 31:1 in the RACE experiments. Since only −1.7- and −1.6-kb promoter sequences for the housekeeping and erythroid promoter, respectively, were analyzed, and since the in vitro expression of a reporter gene does not necessarily reproduce the in vivo situation, it is likely that other regulatory sequences in the URO-synthase gene regulate the activity of these promoters by either enhancing the activity of the erythroid promoter or inhibiting the activity of the housekeeping promoter. For example, while transcription initiation from the murine hydroxymethylbilane synthase housekeeping promoter was high in MEL cells, it was blocked from elongation (43, 45).

The transcription factor consensus binding sites in the erythroid URO-synthase promoter region (Fig. 6) that are significant with respect to erythroid-specific control of transcription include vMYB (nt −27), GATA1 (nt −63, −200, and −383), E47 (nt −146 and −160), a CCAAT box (nt −324), and an NF-E2-like site (nt −400). GATA1 is essential for erythroid development. In chimeric embryos or mice with a GATA1 knockout, erythroid differentiation fails to proceed beyond the proerythroblast stage (46). E47 has been shown to be a DNA binding partner of GATA1 and Tal1 in the formation of a transcriptional transactivating complex in erythroid cells (47) and may play a supporting role in the erythroid-specific control of URO-synthase expression. CCAAT is an enhancer element bound by a family of ubiquitous and tissue-specific factors (e.g. NF-Y (48), which enhance the transcription of many eukaryotic genes. It occurs more frequently in TATA-less promoters (49) and presumably participates in the activation of erythroid URO-synthase transcription. The coincidence of the multiple Inr sites, also found in other TATA-less promoters (50), and the clusters of 5'-RACE ends suggest that these are the functional cap sites for the URO-synthase erythroid transcripts. MYB is a protein target for binding by the murine c-Maf protein, and this complex has been shown to inhibit gene expression in myeloid cells (51). Interestingly, c-Myc knockout mice die in utero from a severe anemia due to an inability to switch from yolk sac to
fetal liver erythropoiesis (52). Finally, NF-E2 is important in the chromatin remodeling process required to prepare a promoter region for transcription. The NF-E2 element in the β-globin locus control region was shown to direct chromatin remodeling in the promoter region of the epsilon-globin gene in a functional minichromosome, activating it for erythroid-specific transcription (53). However, NF-E2 knockout mice were only mildly anemic and hypochromic but had no platelets and died of a bleeding diathesis (54), indicating the primary role of this transcription factor in megakaryocyte maturation. Functional analyses with promoter-reporter constructs in various cell types coupled with footprint and EMSA analyses would facilitate the delineation of transcription factor binding site occupancy for these putative binding sites and for those discussed below for the housekeeping promoter.

Many of the transcription factor binding sites in the URO-synthase housekeeping promoter (Fig. 5) are used by ubiquitously expressed transcription factors (Sp1, NF1, AP1, Oct1, and NFR2), which would provide appropriate basal expression of the housekeeping URO-synthase transcripts in a variety of cell types. For example, NFR2 (nuclear respiratory factor-2) is a ubiquitously expressed activator of various nuclear encoded enzymes in the pathway, have alternative promoters in both species (5, 12, 18, 24, 25). The expression of URO-synthase expression has important implications for the human disorder produces primarily an erythroid porphyria (233, 729–738).

This identification of the erythroid-specific regulation of URO-synthase expression has important implications for the human porphyrias. The use of an alternative erythroid-specific promoter is likely in the human URO-synthase gene, since the gene organization and regulation of the other murine and human heme biosynthetic genes are highly homologous. For example, the first and rate-limiting enzyme, 5-aminolevulinate synthase, is encoded by erythroid-specific and housekeeping transcripts of separate genes (62, 63), the erythroid-specific gene being localized to the X chromosome in mice and humans and the housekeeping gene on chromosome 3p21 in humans but unmapped in mice (64–66). Murine and human 5-aminolevulinate dehydratase and hydroxymethylbilane synthase (also known as porphobilinogen deaminase), the second and third enzymes in the pathway, have alternative promoters in both genes in both species controlling the synthesis of erythroid-specific and housekeeping transcripts (2, 9, 41–44). Similarly, both coproporphyrinogen oxidase and ferrochelatase have been shown to have a single gene and promoter in both species (5, 12, 67, 68). Thus, it is likely that the human URO-synthase gene will be regulated by an alternative erythroid-specific promoter and that studies of its differential expression in erythroid and nonerythroid tissues may clarify why the defective enzyme in the human disorder produces primarily an erythroid porphyria as opposed to a hepatic or erythropoietic porphyria. Furthermore, this mode of regulation may provide insight into why 15% of the CEP patients do not exhibit mutations in their coding sequences. Indeed, we recently identified a CEP patient with severe disease who was heteroallelic for a mutation (T–70 → C) in the first GATA1 binding site in the human URO-synthase erythroid promoter and a coding region mutation (69).

In summary, the mouse URO-synthase gene has been cloned and characterized. Transcription generates three alternative mRNAs differing in their 5′-UTRs; two are expressed by a housekeeping promoter, while the third is expressed by an erythroid-specific promoter, which contains erythroid-specific transcription factor binding elements. Functional assays with promoter-reporter constructs demonstrated that the housekeeping promoter expressed ubiquitous transcripts, whereas the erythroid promoter expressed transcripts only in erythroid cells. Thus, these studies show that the expression of URO-synthase is differentially regulated by transcriptional activation in erythroid cells and potentially provide insight into the molecular pathogenesis of CEP.

Acknowledgments—We thank Dr. Weiming Xu for contributions to this work, Dr. Alicia Alonzo for helpful discussions, and Raman Reddy for expert technical assistance.

REFERENCES
1. Conboy, J. G., Cox, T. C., Bottomley, S. B., Bawden, M. J., and May, B. K. (1992) J. Biol. Chem. 267, 18753–18758
2. Kaya, A. H., Plewiniska, M., Wong, D. M., Desnick, R. J., and Wetmur, J. G. (1994) Genomics 19, 242–248
3. Yoo, H. W., Warner, C. A., Chen, C. H., and Desnick, R. J. (1990) Genomics 15, 21–29
4. Moran-Jimenez, M. J., God, C., Romana, M., Enriquez De Salamanca, R., Teixe, A., Topi, G., D’Alessandro, L., and De Verneuil, H. (1996) Am. J. Hum. Genet. 58, 712–722
5. Delfau-Larue, M. H., Martasek, P., and Grandchamp, B. (1994) Hum. Mol. Genet. 3, 1235–1330
6. Taketani, S., Inazawa, J., Abe, T., Furukawa, T., Kohno, H., Tokunan, R., Nishimura, K., and Inokuchi, H. (1995) Genomics 29, 698–703
7. Taketani, S., Inazawa, J., Nakashashi, Y., Abe, T., and Tokunan, R. (1992) Eur. J. Biochem. 205, 217–222
8. Bishop, D. F., Henderson, J. K., and Astrin, K. A. (1990) Genomics 7, 207–214
9. Grandchamp, B., De Verneuil, H., Beaumont, C., Chretien, S., Walter, O., and Nordmann, Y. (1987) Eur. J. Biochem. 162, 105–110
10. Romana, M., Duhart, A., Runacres, D., Chabret, C., Gossens, M., and Romeo, P. H. (1987) Nucleic Acids Res. 15, 7343–7356
11. Tanabe, A., Furukawa, T., Ogawa, Y., Yamamoto, M., Hayashi, N., Tokunan, R., and Taketani, S. (1997) Biochem. Biophys. Res. Commun. 233, 729–738
12. Tugores, A., Magness, S. T., and Brenner, D. A. (1994) J. Biol. Chem. 269, 30789–30797
13. Bogorad, L. (1958) J. Biol. Chem. 233, 510–518
14. Battersby, A. R., Fooke, C. J. R., Gustafson-Potter, K. E., McDonald, E., and Matcham, G. W. J. (1992) J. Soc. Perkin Trans. 1, 2427–2444
15. Battersby, A. R., Fooke, C. J. R., Gustafson-Potter, K. E., McDonald, E., and Matcham, G. W. J. (1992) J. Soc. Perkin Trans. 1, 2413–2426
16. Tsai, S. F., Bishop, D. F., and Desnick, R. J. (1987) J. Biol. Chem. 262, 1268–1273
17. Kohashi, M., Clement, R. P., Tse, J., and Piper, W. N. (1984) Biochem. J. 220, 755–765
18. Smythe, E., and Williams, D. C. (1988) Biochem. J. 253, 275–279
19. Hart, G. J., and Battersby, A. R. (1985) Biochem. J. 212, 151–160
20. Alwan, A. F., Mege, B. I., and Jordan, P. M. (1989) Biochem. J. 264, 397–402
21. Tsai, S. F., Bishop, D. F., and Desnick, R. J. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 7049–7053
22. W. X., Kozak, C. A., and Desnick, R. J. (1995) Genomics 25, 556–562
23. Shoolingin-Jordan, P. J. (1995) J. Bioenerg. Biomembr. 27, 181–195
24. Romeo, G., and Levin, E. Y. (1968) Science 161, 807–808
25. Levin, E. Y. (1968) Science 161, 807–808
26. Levin, E. Y. (1968) Science 161, 807–808
27. Desnick, R. J., Glass, I. A., Xu, W., Solis, C., and Astrin, K. H. (1998) Semin. Liver Dis. 18, 77–84
28. Bouchechar, S., Da Silva, V., Deybach, J. C., Nordmann, Y., Grandchamp, B., and de Verneuil, H. (1992) Hum. Genet. 88, 320–324
29. Warner, C. A., Yoo, H. W., Roberts, A. G., and Desnick, R. J. (1992) J. Clin. Invest. 89, 693–700
30. Xu, W., Warner, C. A., and Desnick, R. J. (1995) J. Clin. Invest. 95, 905–912
31. Astrin, K. H., Warner, C. A., Yoo, H. W., Goodfellow, P. J., Tsai, S. F., and Desnick, R. J. (1991) Hum. Genet. 87, 18–22
32. Quandt, K., Prech, K., Karczew, W., and Wunderer, E., and Terner, W. (1995) Nucleic Acids Res. 23, 4578–4584
33. Breathnach, R., and Chambon, P. (1981) Annu. Rev. Biochem. 50, 349–383
34. Wei, S. T. (1997) Biochim. Biophys. Acta 1351, 73–88
35. Javahery, R., Khachi, A., Lo, K., Zennie Gregory, B., and Smale, S. T. (1994) Mol. Cell. Biol. 14, 116–127
36. Kappas, A., Sassa, S., Galbraith, R. A., and Nordmann, Y. (1995) in The
Erythroid-specific Expression of Mouse URO-synthase