Promoter and Intronic Sequences of the Human Thiopurine S-Methyltransferase (TPMT) Gene Isolated from a Human Pac1 Genomic Library

Eugene Y. Krynetski,[1] Michael Y. Fessing,[1] Charles R. Yates,[1] Daxi Sun,[1] John D. Schuetz,[1] and William E. Evans[1,2]

Received August 22, 1997; accepted September 15, 1997

Purpose. To isolate and characterize the polymorphic human thiopurine S-methyltransferase (TPMT) gene.

Methods. The human TPMT gene was isolated by PCR screening of a phage artificial chromosome (PAC) library, using exon- and intron-specific primers, then mapped and sequenced.

Results. Two separate PAC1 clones were isolated that contained the same 25 kb gene with 9 exons encompassing the entire TPMT open reading frame. Structural characterization revealed distinct differences when compared to a TPMT gene previously isolated from a chromosome 6-specific human genomic library; the 5′-flanking region (putative promoter) contains 17 additional nucleotides located at position -77 upstream from the transcription start site, in addition to several nucleotide sequence differences, and intron 8 is only 1.6 kb, 5 kb shorter than previously reported. Southern and PCR analyses of genomic DNA from 18 unrelated individuals revealed only the TPMT gene structure corresponding to the PAC clones we isolated. Analysis of the TPMT promoter activity using the 5′-terminal region confirmed transcriptional activity in human HepG2 and CCRF-CEM cells. The 5′-flank is 71% GC rich and does not contain consensus sequences for TATA box or CCAAT elements. FISH analysis demonstrated the presence of the TPMT-homologous sequence on the short arm of chromosome 6 (sublocalized to 6p22).

Conclusions. These findings establish the genomic structure of the human TPMT gene, revealing differences in the promoter and intronic sequences compared to that previously reported and providing a basis for future studies to further elucidate its biological function and regulation.

KEY WORDS: thiopurine S-methyltransferase; gene structure; cloning.

INTRODUCTION

Thiopurine S-methyltransferase (TPMT) is a cytosolic enzyme that catalyzes the S-methylation of aromatic and heterocyclic sulfhydryl compounds, including cytotoxic and immunosuppressive thiopurine medications such as mercaptopurine, azathioprine (a mercaptopurine prodrug) and thioguanine (1). The activity of TPMT exhibits genetic polymorphism in Caucasians and African-Americans, with approximately 90% having high activity, ~10% having intermediate activity due to heterozygosity at the TPMT locus and about 1 in 300 inheriting TPMT-deficiency as an autosomal recessive trait (2–3). Clinical studies have established an inverse correlation between mercaptopurine (or azathioprine) activation to thioguanine nucleotides (TGN) and TPMT activity (4), presumably because patients who lack TPMT activity shunt more drug down the activation pathway. Patients who inherit TPMT-deficiency accumulate erythrocyte TGN concentrations that are 10–15 fold higher than patients with high TPMT activity, if treated with conventional dosages of thiopurines (5–6). Several studies have established that TPMT-deficient patients are at high risk of severe and potentially fatal hematopoietic toxicity (7), unless their thiopurine dosage is decreased by 10–15 fold (6,8,9). The genetic basis for the TPMT polymorphism has recently been elucidated by cDNA cloning of mutant alleles in patients with TPMT-deficiency (10–12).

Identification of inactivating mutations in the TPMT gene has permitted development of genotyping methods for molecular diagnosis of TPMT deficiency, to avoid hematopoietic toxicity in affected patients (13). The presence of a processed TPMT pseudogene (14) mapped to chromosome 18, precluded development of effective PCR-based genotyping assays using primers based on TPMT exon sequences. After the TPMT structural gene was mapped to chromosome 6, a human chromosome 6-specific cosmid library was used to isolate a 30-Kb gene with an open reading frame (ORF) corresponding to the wild-type TPMT sequence (15). Independently a 25 kb human TPMT chromosomal gene was isolated and characterized in our lab, providing the requisite intron sequences for developing genotyping assays for diagnosis of TPMT deficiency (13). Herein we report the structure of the human TPMT gene we isolated from a phage artificial chromosome (PAC) library, revealing substantial differences in the 5′-untranslated region (promoter region), a 5 Kb deletion in intron 8 and several nucleotide differences when compared to the TPMT gene isolated from the chromosome 6-specific cosmid library.

METHODS

Human Subjects

Leukocytes were isolated as previously described from peripheral blood obtained from healthy volunteers and children with acute lymphoblastic leukemia (3), and genomic DNA was isolated by chlorophorm/phenol extractions. TPMT phenotype was assigned on the basis of erythrocyte TPMT activity, as previously described (2–3) Patients with TPMT activity indicative of wild type homozygous genotype (12–26 U/ml pRBC) were chosen for further characterization of the TPMT locus. These studies were approved by the institutional review board for clinical trials at St. Jude Children's Research Hospital, and informed consent was obtained from the participants or their guardians.

[1] Dept. of Pharmaceutical Sciences, St. Jude Children's Research Hospital, 332 N. Lauderdale St., Memphis, Tennessee 38101, and University of Tennessee, Memphis.

[2] To whom correspondence should be addressed. (e-mail william.evans@stjude.org)

Supported in part by the following NIH grants: R37 CA36401, Leukemia Program Project grant CA20180, Cancer Center CORE grant CA21765, by an endowment from the First Tennessee Bank, by a Center of Excellence grant from the State of Tennessee, and by American Lebanese Syrian Associated Charities.

Table 1. PCR Conditions for Amplification of Exons 3–10 (ORF Only)

Round	Exon	Upstream primer #	Downstream primer #	Fragment length, bp	Buffer (Invitrogen)	μl/50 μl react. mix.
I	3	1	86	360	B	10
II		1	132	320	N	5
I	4	112	123	258	L	5
II		112	93	226	P	5
I	5	76	83	369	M	10
II		76	124	323	N	5
I	6	125	53	470	J	10
II		125	43	392	J	5
I	7	45	92	483	B	10
II		90	92	386	N	5
I	8	95	96	217	A	10
II		127	96	190	N	5
I	9	114	97	191	F	10
II		128	97	183	N	5
I	10	129	131	301	A	5
II		130	131	269	N	5

Polymerase Chain Reaction

Sequences of primers and PCR Optimizer buffers (Invitrogen, Carlsbad, CA) for nested PCR used to amplify exons 3–10 from genomic DNA are shown in Table 1. Approximately 100 ng of DNA isolated from leukocytes was used for each PCR reaction. After amplification, the PCR products were purified in 5% non-denaturing polyacrylamide gels and sequenced.

"Long" PCR reactions were performed with 5–10 ng PAC clone plasmid DNA or 1 μg human genomic DNA as a template, using the Expand High Fidelity PCR system (Boehringer Mannheim, Indianapolis, IN) according to the supplier's recommendations (Table 2). The PCR products were cloned into pCRII or pCR2.1 plasmids and sequenced as described below.

DNA Sequencing

Bidirectional sequencing of PCR products and plasmids was performed with an Applied Biosystems DNA sequencer Model 373A and the ABI PRISM dye Terminator Cycle Core Sequencing Kit with AmpliTaq® DNA Polymerase, FS (Perkin Elmer, Foster City, CA), in the Center for Biotechnology at SJCRH.

Table 2. Primers for Long PCR Amplification of TPMT Chromosomal Gene

Fragment	Forward primer #	Reverse primer #
introns 1 + 2	88	87
intron 3	29	50
intron 4	51	28
intron 5	5	43
intron 6	31	34
intron 7	41	9
intron 8 + exon 9 + intron 9	13	36
intron 8	13	77
intron 9	78	36

Genomic DNA Library Screening

The intron 6-specific primer #45 was generated after sequence analysis of intron 6 as described elsewhere (13). A human genomic library obtained from normal male fibroblasts, in pCYPAC-1 (PAC) vector (Genome Systems, St. Louis, MO) (16) was screened by PCR with exon-specific primer #52 and intron-specific primer #45 (see Table 3). Two positive clones (Genome Systems control numbers 7718, plate 2D14 and 8277, plate 279 F13) were identified and isolated. Selected clones were maintained and propagated in *E. coli* DH10B strain, as recommended by the manufacturer.

Fluorescence In Situ Hybridization (FISH)

Fluorescence in situ hybridization was performed with the PAC clone 7718 plasmid DNA, as previously described (1).

TPMT Promoter-Reporter Construction and Transient Expression Experiments

The 5′ UTR of the TPMT gene was subcloned as a *Hind III* fragment spanning (−873) to (+736) (numbered according to the previously reported transcription start site, 15) into the *SmaI* site of the plasmid pGL3-Basic, after the insert ends had been made blunt with Klenow fragment. The resulting plasmid with the 5′UTR in direct orientation to the luciferase (luc) gene designated as pprTPMT, was used in transfection experiments with human hepatoblastoma (HepG2) cells and human T-lineage lymphoblastoid (CCRF-CEM) cells.

Data Analysis

The University of Wisconsin Genetic Computer Group software package was used to analyze sequence information (17).

Table 3. Sequence of Primers Used to Characterize the Human TPMT Gene

Primer location[a]	Primer #	Primer sequence
exon3 (+)	1	GCACGGAAGACATATGCTTGTGAGAC
exon5 (+)	5	CAGAAGAACCAATCACCG
exon8 (−)	9	GAAAGAACACACAGGAGATACTG
exon8 (+)	13	GATACAATGTTTTCCCTCCTG
exon5 (−)	28	ACTGATTTCCACACCAACTA
exon3 (+)	29	TGACATTGAAGAGTACTCGG
exon6 (+)	31	GAGTTCTTCGGGGAACATTTCATTG
exon7 (−)	34	CACCTGGATTAATGGCAACTAATGC
exon10 (−)	36	GTCTCATTTACTTTTCTGTAAGTAGAT
exon7 (+)	41	GATTTGGGATAGAGGAG
intron6 (−)	43	AGAAGACACTGTCTTACTCAC
intron6 (+)	45	ATAACAGAGTGGGGAGGCTGC
exon4 (−)	50	GAAAAAATACCCTCAGTCCACTCTTG
exon4 (+)	51	GCATTTAGATACTTTCCTTAAAGGCA
exon7 (−)	52	ATGTAATACGACTCACTATAACCTGGATTAATGGCAAC
intron6 (−)	53	CAGGGACTTTCAGGAGATGAGCTCTTATCTC
intron4 (+)	76	CTTTGAAACCCTATGAACCTG
exon9 (−)	77	GTGGAACATAAAATGGTGG
exon9 (+)	78	GTCCACCATTTTATGTTCC
intron5 (−)	83	TAAATAGGAACCATCGGACAC
intron3 (−)	86	CTCTGCCAATTAAACTACATC
intron3 (−)	87	GGAGACACAAAAATGTGAAG
exon1 (+)	88	CGGCAACCAGCTGTAAGCGAG
intron6 (+)	90	GCAGACGTGAGATCCTAATACC
intron7 (−)	92	CCCAGGTCTCTGTAGTCAAATC
intron4 (−)	93	CACATCCTGTTAAATCACCC
intron7 (+)	95	TCCTTCCCTGCCTTTTGTCC
intron8 (−)	96	ATCAAGAAACTAGGCAACTG
intron9 (−)	97	AATACAGGCATGAGCCAGCAC
intron3 (+)	112	AATGAAAAGTGTTCACCTACC
intron8 (+)	114	CCCTTCTTTCCTGGACTATTC
intron4 (−)	123	TGTGTATTTCCTGAAAATGG
intron5 (−)	124	CACTGAGAAAAACTTTTGTG
intron5 (+)	125	CTTTGCTTGGCCCTCTTTCC
intron7 (+)	127	CTTTGCCTGTGTAGAGAAATG
intron8 (+)	128	AAGTACACCTACGTCATTGG
intron9 (+)	129	TCAAGTGATCCACCTGCCTC
intron9 (+)	130	TGTTGGGATTACAGGTGTGAG
exon10 (−)	131	CCTCAAAAACATGTCAGTGTG
intron3 (−)	132	ACTCGGGAGACACAAAAATG
intron8 (+)	146	CCAGGTAAAGTTGTTG
intron8 (1)	147	AAACTGGAATTATCTCC

[a](+) sense orientation of the primer; (−) antisense orientation of the primer.

RESULTS

Molecular Cloning and Structural Characterization of TPMT Gene

Intron- and exon-specific PCR primers were used for screening a human PAC library and led to isolation of two recombinant PAC clones # 7718 and #8277. Southern hybridization analysis of these clones with oligonucleotides specific for the 5′-end of TPMT cDNA (primer#88) and to the 3′-end of the TPMT ORF (primer #36) (18,10) indicated that both clones contained the 5′-end of the transcribed region and the 3′-end of the sequence present in the ORF (data not shown). Clone #7718 was subsequently used for structural analysis of the gene.

These PAC clones were used to generate restriction maps of the gene, using both long-range PCR and partial digestion with *Bam HI, Eco RI* and *Hind III* restriction endonucleases (Fig. 1). The restriction maps were similar to that previously published (15) with the following exceptions: there were two additional Eco RI restriction sites in introns 2 and 4, absence of one *Bam HI* and one *Eco RI* site in intron 8, and a 5.3 Kb shorter intron 8 (1.7 vs approximately 7 Kb). Restriction map of the second PAC clone #8277 coincided with the map of clone #7718, thus independently confirming these findings. "Long" PCR was used with the TPMT cDNA specific primers to generate DNA fragments encompassing all intronic sequences (Fig. 1). Sequencing of the terminal regions of inserts revealed that all the TPMT exon-intron boundaries contained the consensus splice donor-acceptor (GT-AG) sequence (data not shown). Nucleotide sequence of the coding regions matched perfectly with that of the TPMT cDNA (18). Subsequently, the conditions

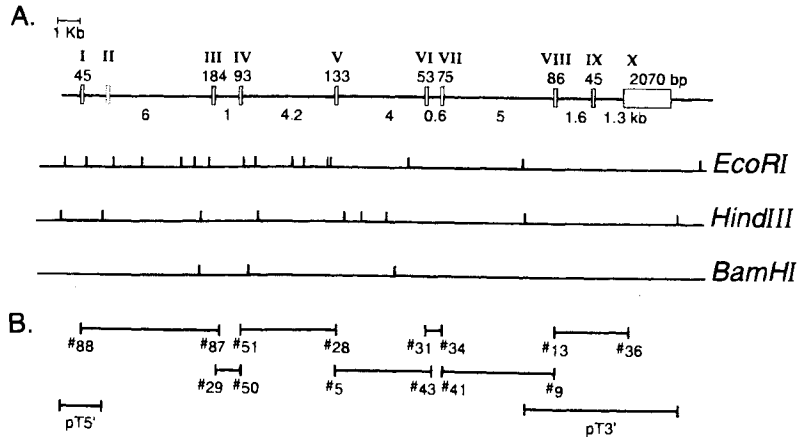

Fig. 1. Human TPMT gene structure. (A) TPMT gene structure and restriction maps. Boxes depict exons in the gene. TPMT exon lengths are shown in bp and the intron lengths are shown in Kb. (B) The position of the intron containing fragments amplified by "long" PCR as described in Materials and Methods. Numeration of the primers and the corresponding sequences are indicated in Table 3. The 5′- and 3′-regions of the gene subcloned into the plasmids pT5′ and pT3′, respectively, are also shown.

for nested PCR were developed for amplification of all the TPMT ORF regions from human genomic DNA, providing a method to identify unknown mutations within the gene ORF and exon-intron boundaries (Table 1). Instead of the "long" PCR technique, 5′- and 3′-termini of the TPMT gene were cloned as *Hind III* fragments using a bacterial vector, by conventional procedures. The resulting plasmids (pT5′ and pT3′) were subsequently used for sequence analysis of 5′- and 3′-regions of the gene (Fig. 1).

No tandemly located genes or pseudogenes were detected in the isolated PAC clones as demonstrated by Southern hybridization. In both cases, only the fragments corresponding to the gene restriction map were found (data not shown). The nucleotide sequences of the TPMT gene reported in this paper have been submitted to the GenBank with accession numbers AF019363-AF019369.

Chromosomal Localization of TPMT Gene

The PAC clone #7718, covering the whole TPMT chromosomal gene (i.e., approximately 25 Kb) was used as a probe to localize the TPMT gene on human lymphocyte metaphase chromosomes, followed by hybridization with a chromosome 6 centromere specific probe. The FISH analysis unambiguously demonstrated the presence of the TPMT-homologous sequence on the short arm of chromosome 6 (sublocalized to 6p22). No signal was found on any other chromosome or any other parts of chromosome 6, consistent with the TPMT gene being present as a single copy in the human genome (1).

Characterization of Intron 8 in Human DNA

To further characterize differences between the previously published TPMT chromosomal gene structure and the TPMT gene we isolated, Southern blot analysis was performed, as well as PCR analysis of DNA isolated from human leukocytes. A DNA probe specific to the 5′-end of intron 8 was generated by amplifying PAC clone 7718 DNA with primers #146 and #147. Total DNA for Southern analysis was completely digested with *Eco RI* and *Bam HI* and hybridized with the PCR product. The anticipated band should be 8 Kb if the "short" intron 8 (without a *Bam HI* or *Eco RI* site, Fig. 1A) is present or 3–4 Kb if the previously reported "long" intron 8 (containing *Bam HI* and *Eco RI* sites) is present (15). The Southern hybridization demonstrated one distinct signal corresponding to a fragment of approximately 8 Kb for all DNA samples from 18 unrelated patients (Fig. 2), indicating the presence of the short intron 8 structure. Southern analysis of PAC clones #7718 and #8277 also revealed an 8 Kb fragment, comparable to that found in genomic DNA samples.

The intron 8-containing fragment of the TPMT gene was also amplified from genomic DNA by "long" PCR, using exon

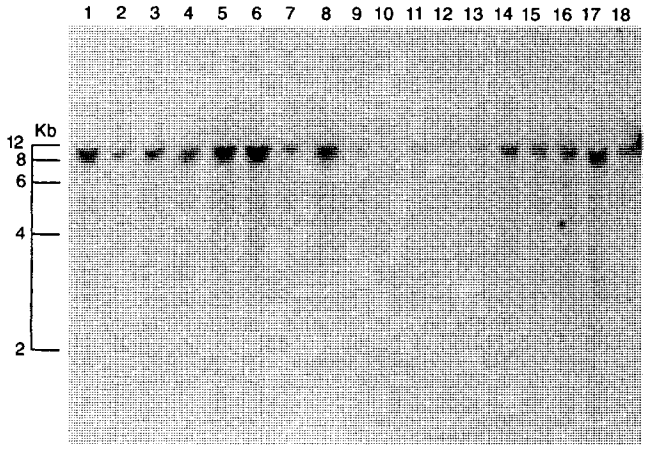

Fig. 2. Southern analysis of the intron 8 region of the human TPMT gene from 18 unrelated individuals (7 females, 11 males; 10 white, 3 black, 3 Oriental and 2 Hispanic; 17 with high TPMT activity [12–26Ushml pRBC] and 1 with intermediate activity [5.2 U/ml pRBC, lane 18]). 5–10 (μg of total DNA from peripheral lymphocytes were completely digested with *BamHI* and *EcoRI* together and subjected to Southern analysis with a probe derived from the 5′ region of intron 8. The probe was generated by PCR amplification of PAC #7718 DNA with primers 146 and 147 (Table 3).

8-specific primer #13 and exon 10- specific primer #36. The amplification with all DNA samples resulted in synthesis of a single fragment of 3.2 Kb (Fig. 3A), corresponding to the structure of the gene we isolated (Fig. 1). No longer fragment was amplified in any samples. The authenticity of the amplified PCR fragments was thereafter confirmed by Southern hybridization: the exon 9-derived PCR product (Table 1) was used as a probe to demonstrate amplification of the DNA region encompassing exon 9 (Fig. 3B). These results indicate that the structure of the intron 8-containing region found in the PAC library was present in genomic DNA of all individuals tested.

5′-Untranscribed Region of TPMT Gene and Its Promoter Activity

Hind III—fragment encompassing 5′-untranscribed region of the TPMT gene, exon 1 and most of intron 1 was cloned in pGEM7Z plasmid (plasmid pT5′ in the Fig. 1B) and sequenced. The region immediately upstream of the previously reported start site (15) showed nucleotide differences and an additional 17-base sequence when compared to the published data (Fig. 4). This sequence contains two putative Sp1 and one KROX-24 (Egr1) binding sites and therefore may be important in transcription activity of the TPMT promoter.

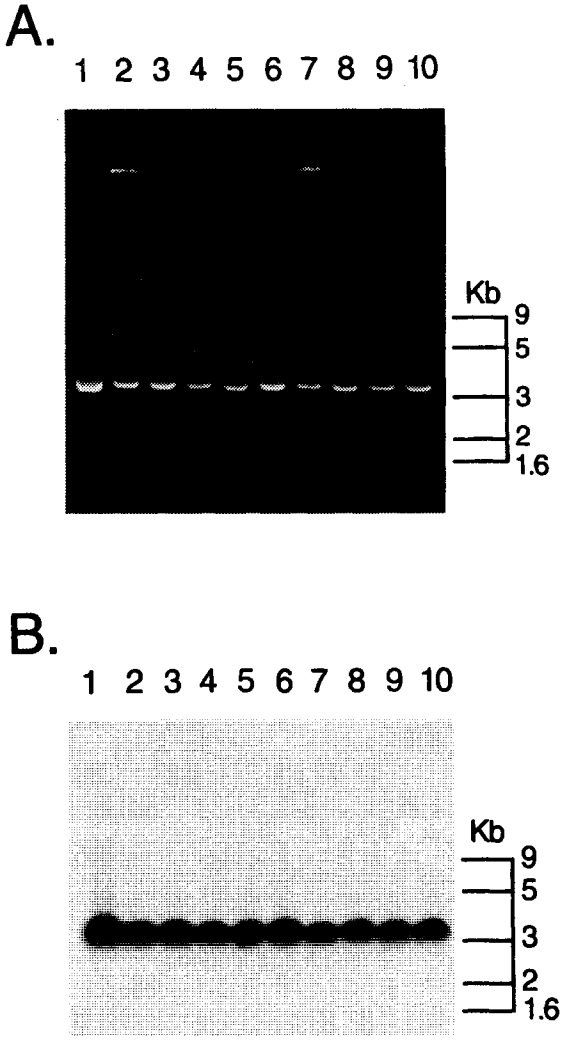

Fig. 3. PCR analysis of the intron 8-containing region of the TPMT gene from 10 unrelated individuals (patients 1–9 correspond to patients 10, 9, 12, 13, 11, 15, 14, 16, 17 in Fig. 2, respectively. Patient 10 in Fig. 3 is not shown in Fig. 2). (A) Ethidium bromide staining of the agarose gel with the PCR products. (B) Southern analysis of the above gel with exon 9-specific probe.

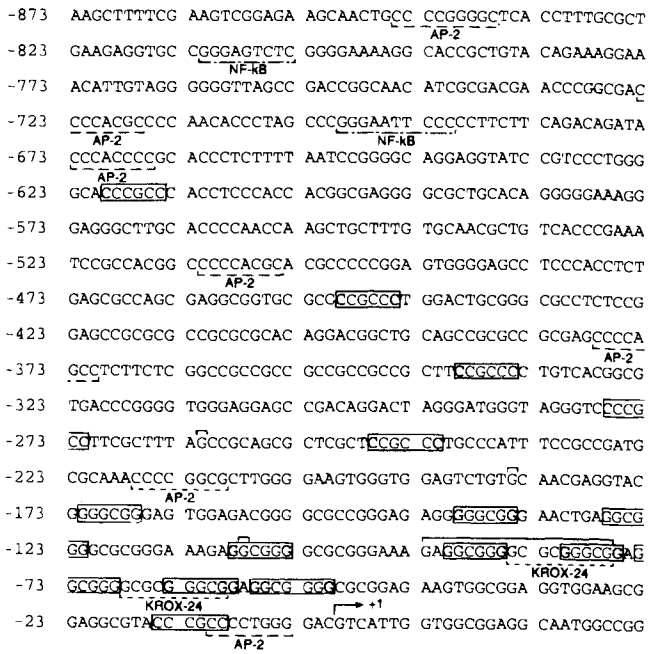

Fig. 4. Nucleotide sequence of the TPMT 5′-flanking region. 17-bp insertion and 3 nucleotide substitutions are marked by a line above the involved sequence. The putative binding sites for the following transcription factors are shown: Sp1 (boxed), NF-kB (underlined with dashed line with dots), AP-2 and KROX-24 (underlined with dashed lines). Transcription start site (position +1) is indicated as reported by Szumlanski et al., 1996.

To assess promoter activity of the 5′-untranscribed region of the TPMT gene, we cloned a *Hind III* fragment spanning (−873) to (+736) into the promoterless luciferase vector (pGL-3Basic). This plasmid pprTPMT with TPMT gene fragment containing the 5′-flank, exon 1 and the 5′ portion of intron 1 was transiently transfected into HepG2 and CCRF-CEM cells and showed 100-fold greater luciferase activity compared to the promoterless pGL-3Basic plasmid. Luciferase activity with pprTPMT was about 10% luciferase activity when compared to pGL-3Control plasmid with luc gene under the control of SV40 early promoter with SV40 enhancer.

DISCUSSION

The TPMT chromosomal gene isolated from a PAC1 human genomic library consists of 10 exons and 9 introns, and spans 25 Kb. This gene encompasses the entire TPMT open reading frame and is identical in exon sequence and structure compared to a TPMT gene previously isolated from a chromosome 6-specific human genetic library (15). However, the gene we isolated differs in size and intron structure and contains an additional 17 nucleotides in the promoter region.

One difference is the length of intron 8, which is 1.7 Kb in our two clones, compared to approximately 7 Kb in that

previously reported. Consistent with this finding, we could not detect the previously reported *Bam HI* and *Eco RI* sites within this intron. The overall distance between the 5′-end of exon 1 and the 3′-end of exon 10 in the previously published map is about 30 Kb, compared to 25 Kb in our clones, coinciding perfectly with the 5 Kb difference in intron 8 in these two clones. Though this difference is unlikely to alter function or regulation pattern of the TPMT gene, it is important for designing PCR conditions for exon 9 amplification to screen for novel mutant alleles at the TPMT gene locus. Southern hybridization with an intron 8-specific probe confirmed the presence of the short variant of intron 8 in genomic DNA from 18 of 18 unrelated patients. This finding was further corroborated using "long" PCR with an exon 8-specific sense primer and an exon 10-specific antisense primer, using genomic DNA from 10 unrelated individuals as templates, and hybridization with an exon 9-specific probe. None of these experiments provided evidence for the longer 7 Kb variant of intron 8. Taken together, these results indicate that the 1.7 Kb intron 8-containing TPMT gene we cloned represents the correct structure of the human TPMT gene, since it was found in two independent PAC clones and in all human DNA samples tested by two different methods (Southern hybridization and "long" PCR). Failure to find the longer intron 8 for this region of the gene as reported by Szumlanski et al. (15) can be explained in two ways: (i) it exists as a rather rare allele of the TPMT gene; (ii) it represents a cloning artifact in the previous report due to some recombination event generated in the human chromosome 6-specific Lawrist4 cosmid library. Other differences include additional *Eco RI* sites found in intron 2 and 4, several minor nucleotide differences in intron sequences and the 17 bp longer promoter region.

Results of the sequencing experiments of PCR products derived from intron-specific primers and encompassing all exons of the open reading frame (i.e., exons 3–9) showed full concordance of exons with corresponding regions of the TPMT cDNA we previously isolated (10). Splice junction sites found in the sequencing experiments follow the "GT-AG" rule, and contain polypyrimidine tracts preceding the 3′-end of introns, consistent with their role in RNA maturation (19). The existence of exon 2 has been observed in one of 16 human liver cDNA's (15). This is not unexpected, as deletion variants of TPMT cDNA had been found in our experiments in two unrelated patients with TPMT activity indicative of a heterozygous genotype (8.3 and 5.2 U/ml of pRBC). These deletions coincided perfectly with boundaries of exon 5–exon 7 (exon 6 deleted) in one clone and exon 8–exon 10 in the other (exon 9 deleted). In two unrelated patients, we found an anomalously spliced RNA where a cryptic splice site was used instead of a normal donor site at the 5′-end of intron 3 (Fig. 5).

Transient transfection of the TPMT 5′-region fused to a luciferase reporter gene demonstrated functional activity of the putative promoter, providing further evidence that we have cloned the functionally active TPMT gene. Computer analysis of the promoter region revealed that it is very GC rich (71% in the region between (−873) and (+1) positions) and lacks a TATA-box, typical features for "housekeeping genes" (e.g., DHFR). By analogy with these genes, the transcription factor Sp1 is probably important for TPMT activity (20). Binding sites for several other transcription factors are also evident in this region, including NF-kB (21), AP-2 (22), and KROX-24

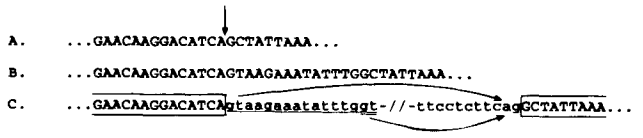

Fig. 5. Comparison of the sequence of wild type TPMT cDNA (A) and the corresponding region of TPMT cDNA isolated and cloned from two heterozygous patients (only a part of the sequence is shown) (B). The relevant fragments of the TPMT chromosomal gene, including exon 3—intron 3 and intron 3—exon 4 junctions, are also shown (C). The boundary between exon 3 and exon 4 in a wild-type TPMT cDNA is indicated with a vertical arrow. Exon and cDNA sequences are in uppercase, and intron sequence in lowercase. The alternative donor site within the intron 3 is double-underlined, and the intron sequence found in the alternatively spliced cDNA is underlined.

(23), as shown in Fig. 4. Because TPMT is expressed in many tissues, it seems unlikely that tissue specific transcription factor binding sites are important in its regulation, although this remains unknown. The TPMT gene we cloned has several nucleotide differences and an additional 17-base sequence found in the 5′-terminus encompassing exon 1 and the adjacent upstream sequence (Fig. 4), which are absent in the previously published variant of the TPMT gene. Comparison of the 5′ promoter region we cloned to a promoter region with the 17-base deletion as reported by Szumlanski et al (15), revealed similar luciferase activity when transfected into HepG2 cells. However, this 17 base fragment contains one KROX-24 and two Sp1 binding sites (Fig. 4), and their importance for TPMT gene regulation remains unknown. Ongoing studies of TPMT gene regulation should elucidate the functional significance of these sequence differences, and define the cis- and trans- elements of the TPMT promoter.

Four lines of evidence support the fact that we have isolated the human gene encoding authentic TPMT: first, the sequence analysis showed full concordance between the exons and the TPMT cDNA (enzymatic activity of the latter was previously confirmed by heterologous expression in several expression systems; (18,10–12)); second, splice junction sites follow the consensus sequence and contain elements necessary for RNA processing; third, the promoter directs abundant transcription in liver cells, one of the tissues with the highest level of TPMT expression; and fourth, its chromosomal localization coincides with that detected by an independent method (PCR analysis of a panel of deletion mutants in chimeric human/rodent cells, (15)). Further evidence for authenticity of the isolated gene comes from genotyping experiments of patients with impaired TPMT activity: genotyping assays developed using intron sequence information from our clones had approximately 95% concordance with TPMT phenotype in a Caucasian population (13). Having isolated the TPMT chromosomal gene and defined its structure, it will now be possible to elucidate the molecular mechanisms of TPMT gene regulation.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the excellent technical assistance of Amy E. Atkinson, and the expertise of Dr. Clayton Naeve in the SJCRH Center for Biotechnology.

REFERENCES

1. E. Y. Krynetski, H. L. Tai, C. R. Yates, M. Y. Fessing, T. Loennechen, J. D. Schuetz, M. V. Relling, and W. E. Evans. *Pharmacogenetics* **6**:279–290 (1996).
2. R. M. Weinshilboum, and S. L. Sladek. *Am. J. Hum. Genet.* **32**:651–662 (1980).
3. H. L. McLeod, J. S. Lin, E. P. Scott, C. H. Pui, and W. E. Evans. *Clin. Pharmacol. Ther.* **55**:15–20 (1994).
4. L. Lennard, J. S. Lilleyman, J. Van Loon, and R. M. Weinshilboum. *Lancet* **336**:225–229 (1990).
5. L. Lennard, J. A. Van Loon, J. S. Lilleyman, and R. M. Weinshilboum. *Clin. Pharmacol. Ther.* **41**:18–25 (1987).
6. W. E. Evans, M. Horner, Y. Q. Chu, D. Kalwinsky, and W. M. Roberts. *J. Pediatr.* **119**:985–989 (1991).
7. E. Schutz, J. Gummert, F. Mohr, and M. Oellerich. *Lancet* **341**:436 (1993).
8. H. L. McLeod, D. R. Miller, and W. E. Evans. *Lancet* **341**:1151 (1993).
9. L. Lennard, B. E. Gibson, T. Nicole, and J. S. Lilleyman. *Arch. Dis. Child* **69**:577–579 (1993).
10. E. Y. Krynetski, J. D. Schuetz, A. J. Galpin, C. H. Pui, M. V. Relling, and W. E. Evans. *Proc. Natl. Acad. Sci. USA* **92**:949–953 (1995).
11. H. L. Tai, E. Y. Krynetski, C. R. Yates, T. Loennechen, M. Y. Fessing, N. F. Krynetskaia, and W. E. Evans. *Am. J. Hum. Genet.* **58**:694–702 (1996).
12. H. L. Tai, E. Y. Krynetski, E. G. Schuetz, Y. Yanishevski, and W. E. Evans. *Proc. Natl. Acad. Sci. USA* **94**:6444–6449 (1997).
13. C. R. Yates, E. Y. Krynetski, T. Loennechen, M. Y. Fessing, H. L. Tai, C. H. Pui, M. V. Relling, and W. E. Evans. *Ann. Int. Med.* **126**:608–614 (1997).
14. D. Lee, C. Szumlanski, J. Houtman, R. Honchel, K. Rojas, J. Overhauser, E. D. Wieben, and R. M. Weinshilboum. *Drug Metab. Dispos.* **23**:398–405 (1995).
15. C. Szumlanski, D. Otterness, C. Her, D. Lee, B. Brandriff, D. Kelsell, N. Spurr, L. Lennard, E. Wieben, and R. Weinshilboum. *DNA Cell Biol.* **15**:17–30 (1996).
16. P. A. Ioannou, C. T. Amemiya, J. Garnes, P. M. Kroisel, H. Shizuya, C. Chen, M. A. Batzer, and P. J. de Jong. *Nat. Genet.* **6**:84–89 (1994).
17. Genetics Computer Group: Program manual for the GCG Package, version 7. April (1991).
18. R. Honchel, I. A. Aksoy, C. Szumlanski, T. C. Wood, D.M. Otterness, E. D. Wieben, and R. M. Weinshilboum. *Mol. Pharmacol.* **43**:878–887 (1993).
19. P. Senapathy, M. B. Shapiro, and N. L. Harris. *Methods Enzymol.* **183**:252–278 (1990).
20. J. T. Kadonaga, K. A. Jones, and R. Tjian. *TIBS* **11**:20–23 (1986).
21. M. J. Lenardo, and D. Baltimore, *Cell* **58**:227–229 (1989).
22. T. Williams, and R. Tjian. *Genes Develop.* **5**:670–682 (1991).
23. B. Christy, and D. Nathans. *Proc. Natl. Acad. Sci. USA* **86**:8737–8741 (1989).