THIOPURINE PHARMACOGENETICS: CLINICAL AND MOLECULAR STUDIES OF THIOPURINE METHYLTRANSFERASE

RICHARD WEINSHILBOUM

Department of Pharmacology, Mayo Medical School/Mayo Graduate School/Mayo Clinic, Rochester, Minnesota

This paper is available online at http://dmd.aspetjournals.org

ABSTRACT:

Thiopurine drugs are used to treat patients with neoplasia and autoimmune disease as well as transplant recipients. These agents are metabolized, in part, by S-methylation catalyzed by thiopurine methyltransferase (TPMT). The discovery nearly two decades ago that levels of TPMT activity in human tissues are controlled by a common genetic polymorphism led to one of the best examples of the potential importance of pharmacogenomics for clinical medicine. Specifically, it is now known that patients with inherited very low levels of TPMT activity are at greatly increased risk for thiopurine-induced toxicity such as myelosuppression when treated with standard doses of these drugs, while subjects with very high activity may be undertreated. Furthermore, recent reports indicate that TPMT may be the target for clinically significant drug interactions and that this common genetic polymorphism might be a risk factor for the occurrence of therapy-dependent secondary leukemia. In parallel with these clinical reports, the molecular basis for the TPMT polymorphism has been determined as a result of cloning and characterization of the human TPMT cDNA and gene. Those advances led to the description and characterization of a series of single nucleotide polymorphisms that result in low levels of enzyme activity as well as a polymorphic variable number tandem repeat within the 5′-flanking region of the TPMT gene that may “modulate” level of enzyme activity. As a result of these observations, the TPMT genetic polymorphism represents a model system for the way in which basic pharmacogenetic information is developed and applied to clinical medicine.

6-Mercaptopurine, 6-thioguanine, and azathioprine are thiopurine drugs that are used to treat acute lymphoblastic leukemia, autoimmune disorders, inflammatory bowel disease and organ transplant recipients (Paterson and Tidd, 1975; Lennard, 1992). Thiopurines are very useful drugs, but they have a relatively narrow therapeutic index, with life-threatening myelosuppression as a major toxicity (Paterson and Tidd, 1975; Lennard, 1992). These drugs are metabolized, in part, by S-methylation catalyzed by the cytoplasmic S-adenosyl-L-methionine-dependent enzyme thiopurine methyltransferase (TPMT), EC 2.1.1.67 (Remy, 1963; Woodson and Weinshilboum, 1983). The level of TPMT activity in human tissues is controlled by a common genetic polymorphism (Weinshilboum and Sladek, 1980), and this polymorphism provides one of the better examples of the potential impact of pharmacogenomics on clinical medicine (Weinshilboum et al., 1999). Therefore, the TPMT polymorphism is a model system that illustrates the process by which pharmacogenetic information is transferred into the clinic as well as the migration over the past two decades of pharmacogenetic research from studies of phenotype to include studies of genotype. The subsequent discussion will briefly review the way in which biochemical pharmacogenetic understanding of TPMT was “translated” into clinical laboratory-based testing that has helped make it possible to individualize thiopurine therapy. In addition, the parallel process by which molecular understanding of the TPMT polymorphism has taken place will be described. Finally, unanswered questions with regard to this area of pharmacogenetic research will also be summarized.

TPMT: Clinical Studies

S-Methylation was reported to be a metabolic pathway for thiopurine drugs soon after these agents were initially administered to humans (Elion, 1967). The first studies of the enzyme that catalyzed this reaction were performed by Remy in the early 1960s using rodent tissue (Remy, 1963). However, it was not until the late 1970s that TPMT activity was first assayed and studied in human tissue (Weinshilboum et al., 1978), with the clearly stated goal of testing the hypothesis that individual variation in this pathway for thiopurine biotransformation might be related to individual differences in drug toxicity and/or therapeutic efficacy. Those initial TPMT enzyme assays in humans were performed with an easily accessible tissue, the red blood cell (RBC). That was done to make it possible to develop a clinical test if it could be demonstrated that RBC TPMT activity was correlated with enzyme activity at sites of thiopurine metabolism such as the liver.

The first application of measurements of RBC TPMT activity involved pharmacogenetic experiments performed with large population samples and nuclear families (Weinshilboum and Sladek, 1980). Those studies demonstrated that the “trait” of level of RBC TPMT activity was controlled by a common genetic polymorphism. Approximately 89% of the Caucasian population studied was homozygous for a gene or genes that resulted in high activity; 11% were heterozygous and had intermediate activity; and 1 of every 300 subjects was...
homzygous for the trait of low level of RBC TPMT activity (Fig. 1). In retrospect, it is fortunate that the initial pharmacogenetic studies of TPMT were performed with Caucasian subjects. Had they been conducted in East Asia—as described subsequently—the trimodal frequency distribution shown in Fig. 1 would not have been observed (Jang et al., 1996; Park-Hah et al., 1996). After the initial population and family-based pharmacogenetic study of RBC TPMT was reported in 1980, work on the polymorphism during the remainder of the decade focused on experiments designed to characterize the biochemical properties of the enzyme (Woodson and Weinshilboum, 1983; Woodson et al., 1983; Ames et al., 1986) as well as experiments demonstrating that genetically determined levels of TPMT activity in the RBC reflected relative levels of TPMT enzymatic activity and immunoreactive protein in other human tissues and cells such as kidney, liver, and lymphocyte (Szumlanski et al., 1992; Van Loon and Weinshilboum, 1982; Woodson et al., 1982).

The steady increase in knowledge of the biochemistry and regulation of TPMT in human tissues during the 1980s was paralleled by increasing understanding of the biotransformation of thiopurine drugs in humans. It was demonstrated that thiopurines such as 6-mercaptopurine are prodrugs that undergo metabolic activation to form 6-thioguanine nucleotides (6-TGNs) (Lennard and Maddocks, 1983). Levels of 6-TGNs measured in the RBC were correlated with both thiopurine therapeutic efficacy and toxicity such as myelosuppression (Lennard et al., 1983). However, the question of why patients treated with similar or identical doses of thiopurine drugs might have very different RBC 6-TGN concentrations remained unanswered. The convergence of research on the pharmacogenetics of TPMT with thiopurine clinical pharmacology resulted in the understanding that, although thiopurines are substrates for several metabolic pathways (see Fig. 2), a major factor responsible for individual variation in 6-TGN concentrations was the genetically determined level of TPMT activity (Lennard et al., 1987, 1990). Specifically, the higher the level of TPMT activity, the lower the 6-TGN levels in the RBC, and vice versa (Fig. 3). It was a short step from those observations to the first report in 1989 that patients with genetically very low or absent TPMT activity who were treated with “standard” doses of these drugs were at risk for the development of life-threatening thiopurine toxicity (Lennard et al., 1989). Furthermore, evidence was obtained that patients with very high TPMT activity might display decreased therapeutic efficacy when treated with standard doses of thiopurine drugs (Lennard et al., 1990). Those same clinical studies also raised the possibility that the level of TPMT activity might be induced during chronic drug therapy. Specifically, it was observed that the level of RBC enzyme activity decreased an average of approximately 25% after drug therapy of leukemia was terminated (Lennard et al., 1990), suggesting that the enzyme activity was increased during chronic therapy. That issue will be discussed again under TPMT: Outstanding Issues. Taken together, these developments offered the prospect of an enhanced ability to individualize thiopurine drug therapy. Since those early clinical reports, the association of very low TPMT activity with thiopurine toxicity has been confirmed repeatedly (Weinshilboum et al., 1999). Consequently, in some referral centers the measurement of TPMT enzyme activity has become a standard clinical test (Mayo Foundation Laboratory Medicine Bulletin, 1991), one of the first examples of a pharmacogenetic test that has entered the mainstream of clinical practice. However, the time required for the transfer of this pharmacogenetic information into medical practice should be noted. The initial observation that genetically low TPMT activity was associated with an increased risk of drug toxicity appeared in 1989 (Lennard et al., 1989) and was subsequently confirmed by a series of clinical reports that appeared throughout the 1990s. Those reports involved disorders as diverse as acute lymphoblastic leukemia, dermatologic

---

**Fig. 1.** RBC TPMT frequency distribution histogram for 298 randomly selected Caucasian subjects.

Modified from Weinshilboum and Sladek (1980). Reproduced with permission of the University of Chicago Press.

**Fig. 2.** Thiopurine biotransformation.

A simplified outline of the biotransformation of azathioprine and 6-mercaptopurine to yield 6-TGN. XO, xanthine oxidase; HPRT, hypoxanthine phosphoribosyl transferase; IMP, inosine monophosphate; and GMP, guanosine monophosphate.

**Fig. 3.** Relationship between RBC TPMT activity and RBC 6-TGN concentrations in 95 children with acute lymphoblastic leukemia (ALL) who were treated on protocol UK ALL VIII.

Modified from Lennard et al. (1990). Reproduced with permission of The Lancet.
disease, renal and cardiac transplantation, rheumatoid arthritis, and autoimmune hepatitis (Evans et al., 1991; Anstey et al., 1992; Schütz et al., 1993; Kerstens et al., 1995; Ari et al., 1995; Escouesse et al., 1995). The number of these clinical reports increased gradually through the 1990s, peaking in the middle of the decade. Therefore, the process of clinical validation and confirmation required nearly a decade after the first report of an association between genetically decreased TPMT activity and life-threatening thiopurine-induced myelosuppression. It will be of interest to compare this time course with that of the translation processes for pharmacogenetic information in the future—at a time when advances in DNA-based technology may make it possible to provide clinicians with pharmacogenetic data in a much more rapid and cost-effective fashion.

Understanding the role of TPMT in thiopurine biotransformation has provided insights that extend beyond pharmacogenetics to include the possibility of clinically significant drug interactions. During the biochemical characterization of TPMT in the early 1980s, it was noted that benzoic acid derivatives such as salicylic acid were potent inhibitors of the enzyme (Woodson et al., 1983; Ames et al., 1986). Those observations raised the possibility of drug interactions if TPMT proved, as it has, to play an important role in individual variation in thiopurine toxicity and/or therapeutic efficacy. However, it required over a decade for the first reports to appear of possible interactions between thiopurines and TPMT inhibitors such as the aminosalicylic acid derivatives that are used to treat inflammatory bowel disease (Griffin and Miner, 1995; Hanauer, 1996). Sulfasalazine and other aminosalicylic acid derivatives used to treat these patients were shown to be potent in vitro inhibitors of recombinant human TPMT (Szumlanski and Weinshilboum, 1995), and there has now been at least one report of a potentially serious drug interaction when these agents were administered to a patient who was also being treated with standard doses of thiopurine drugs (Lewis et al., 1997). That report raised the possibility that such patients can be converted to “phenocopies” of subjects with genetically low TPMT activity.

Finally, the clinical implications of the TPMT polymorphism may not end with acute toxicity, variations in therapeutic efficacy, or drug interactions. It has been known for some time that treatment with antineoplastic agents can be associated with the occurrence of late secondary neoplasia, which is thought to result from the initial chemotherapy. For example, children with acute lymphoblastic leukemia who are treated with drug regimens that include thiopurines and are initially “cured” can later develop acute myelogenous leukemia that is thought to be treatment related (Ratain and Rowley, 1992). Although many factors are involved in the occurrence of this tragic outcome of pharmacologic therapy, several reports have now indicated that decreased TPMT activity may be one risk factor for the occurrence of either secondary myelodysplastic syndrome or acute myelogenous leukemia in these patients (Relling et al., 1998; Thomsen et al., 1999).

In summary, the TPMT genetic polymorphism represents an excellent example of the application of pharmacogenetics in the clinic. Subjects with genetically low TPMT activity are at greatly increased risk for thiopurine-induced toxicity, and those with very high TPMT activity may require treatment with slightly elevated doses of these drugs—always with careful clinical monitoring. Furthermore, caution should be exercised when patients are treated simultaneously with thiopurine drugs and agents such as aminosalicylic acid derivatives that are known to inhibit TPMT. Finally, the possibility that decreased TPMT activity might represent a risk factor for the occurrence of late secondary neoplasia in patients treated with thiopurines deserves further study. In parallel with, and complementary to, this rapidly evolving understanding of the clinical relevance of the TPMT genetic polymorphism has been a steady increase in knowledge of the molecular basis for the genetic regulation of TPMT.

TPMT: Molecular Studies

Understanding the molecular basis for inherited variation in drug response has been a major goal of pharmacogenetic research. Therefore, in parallel with the clinical studies described in the preceding paragraphs, steady progress has also occurred as a result of research into molecular mechanisms responsible for phenotypic differences in the level of TPMT activity. That understanding resulted from the application of a “classical” molecular biology research strategy. As the initial step, human kidney TPMT was purified and photoaffinity labeled, and partial amino acid sequence was obtained (Van Loon et al., 1992). That information made it possible to successfully clone a human TPMT cDNA (Honchel et al., 1993). Northern blot analysis performed with the TPMT cDNA as a probe showed that mRNA for the enzyme was widely expressed in many human tissues (Lee et al., 1995), raising the possibility of a function for TPMT that extends beyond the biotransformation of xenobiotics such as thiopurine drugs. That issue will be raised again under TPMT: Outstanding Issues. However, at the time that the cDNA was cloned, the immediate goal was to determine the molecular basis for the common genetic polymorphism that regulated the trait of the level of enzyme activity. Achieving that goal would be greatly simplified by knowledge of the TPMT gene structure. However, attempts to use the human TPMT cDNA to clone the gene were complicated by the presence of a processed pseudogene for TPMT that was located on the long arm of chromosome 18 (Lee et al., 1995). The existence of the processed pseudogene is not merely of intellectual interest, since it is important that investigators using reverse transcription-polymerase chain reaction (RT-PCR) to study TPMT avoid inadvertently amplifying this intronless pseudogene—thus mistaking it for a variant mRNA species—as a result of DNA contamination of their RNA preparations. It is also important that PCR primers used to perform amplifications of genomic DNA include at least one intron-based primer to avoid amplifications of the TPMT-processed pseudogene. Subsequently, as a result of the use of anchored PCR techniques that made it possible to map the active gene to chromosome 6, the gene was cloned from a chromosome 6-specific cosmid library, thus avoiding the processed pseudogene on chromosome 18 (Szumlanski et al., 1996). TPMT is approximately 34 kilobases in length, consists of 10 exons—8 of which encode protein—and maps to chromosome 6p22.3 (Szumlanski et al., 1996).

The article describing the TPMT gene structure also reported two single nucleotide polymorphisms (SNPs) present in the most common variant allele for low TPMT activity in Caucasians—TPMT*3A (Fig. 4) (Szumlanski et al., 1996). Both of these SNPs, one located in exon 7 and the other in exon 10, altered the encoded amino acid, i.e., they were nonsynonymous cSNPs. That study also demonstrated that the presence of these two SNPs was associated with a decrease in levels of immunoreactive TPMT protein during transient expression in COS-1 cells (Szumlanski et al., 1996). It was reported later that this effect was primarily due to an increased rate of degradation of the variant TPMT protein encoded by this allele—probably through a proteasome-mediated mechanism (Tai et al., 1997). A series of variant TPMT alleles has now been described, virtually all involving SNPs present in the most common intronless pseudogene—thus mistaking it for a variant mRNA species—as a result of DNA contamination of their RNA preparations. It is also important that PCR primers used to perform amplifications of genomic DNA include at least one intron-based primer to avoid amplifications of the TPMT-processed pseudogene. Subsequently, as a result of the use of anchored PCR techniques that made it possible to map the active gene to chromosome 6, the gene was cloned from a chromosome 6-specific cosmid library, thus avoiding the processed pseudogene on chromosome 18 (Szumlanski et al., 1996). TPMT is approximately 34 kilobases in length, consists of 10 exons—8 of which encode protein—and maps to chromosome 6p22.3 (Szumlanski et al., 1996).

The article describing the TPMT gene structure also reported two single nucleotide polymorphisms (SNPs) present in the most common variant allele for low TPMT activity in Caucasians—TPMT*3A (Fig. 4) (Szumlanski et al., 1996). Both of these SNPs, one located in exon 7 and the other in exon 10, altered the encoded amino acid, i.e., they were nonsynonymous cSNPs. That study also demonstrated that the presence of these two SNPs was associated with a decrease in levels of immunoreactive TPMT protein during transient expression in COS-1 cells (Szumlanski et al., 1996). It was reported later that this effect was primarily due to an increased rate of degradation of the variant TPMT protein encoded by this allele—probably through a proteasome-mediated mechanism (Tai et al., 1997). A series of variant TPMT alleles has now been described, virtually all involving SNPs within the cDNA open reading frame (ORF) that result in missense or nonsense codons (Krynetski et al., 1995; Otterness et al., 1997). However, one SNP at the TPMT intron 9/exon 10 splice junction that disrupted the canonical sequence found at splice sites segregated with a decreased level of enzyme activity within an extended kindred (Otterness et al., 1998). This polymorphism resulted in the activation
of a cryptic splice site within intron 9 as well as a significant decrease in the level of TPMT mRNA (Otterness et al., 1998).

As mentioned earlier, population studies showed that the frequency distribution of RBC TPMT activity in East Asians did not display the trimodal frequency distribution shown in Fig. 1 (Jang et al., 1996; Park-Hah et al., 1996). Subsequent molecular epidemiology studies indicated that the most common TPMT variant allele in Caucasians, TPMT*3A, is either not present in Asian populations or is present at a very low frequency (Otterness et al., 1997; Collie-Duguid et al., 1998). In those groups, the most common variant allele is TPMT*3C, an allele that includes only the exon 10 SNP (Fig. 4). Although the majority of the presently known variant alleles that alter TPMT activity are due to SNPs within the ORF, even subjects with the trait of “high” TPMT activity show a wide range in the level of their enzyme activity (see Fig. 1), and that variance has been reported to result predominantly from the effects of inheritance (Vuclietich et al., 1995). Those observations raised the possibility that additional molecular genetic mechanisms beyond SNPs within the ORF might be involved in the regulation of the level of TPMT activity. A search for such mechanisms led to the discovery of functional polymorphisms within the 5' flanking region of TPMT.

When the structure of the human TPMT gene was reported, it was noted that it did not have a canonical TATA box sequence located near the site of transcription initiation (Szymanski et al., 1996). However, the 5’-flanking region of the gene was GC-rich, with a series of potential Sp1 binding sites (Szymanski et al., 1996). Those observations were pursued systematically by Spire-Vayron de la Moureyre et al. (1998, 1999) in a series of studies demonstrating that the 5'-flanking region of TPMT included a polymorphic variable number tandem repeat (VNTR) in which a 17- or 18-base pair repeat element was present from four to eight times. The most common VNTR alleles in the original Caucasian population sample studied had four or five repeat elements (Spire-Vayron de la Moureyre et al., 1998, 1999). A subsequent, larger study confirmed that TPMT VNTR*4 and *5 were the most common alleles in Caucasians, with repeat numbers that varied from three to nine (Yan et al., 2000). Most interestingly, there was an inverse relationship between the sum of repeat units (i.e., the number of repeats on the two alleles added together) and level of RBC TPMT activity (Spire-Vayron de la Moureyre et al., 1999; Yan et al., 2000). Those empirical observations were confirmed and extended by studies conducted with reporter gene constructs that showed a decrease in the level of reporter gene expression as the number of repeat elements increased (Spire-Vayron de la Moureyre et al., 1999). Whether this recently observed mechanism for modulating the level of TPMT activity is of potential clinical importance remains to be determined.

TPMT: Outstanding Issues

Although we have learned a great deal about TPMT during the past two decades—both with regard to clinical pharmacogenetics and molecular genetic mechanisms involved in its regulation—many unanswered questions remain. The first of these relates to the possibility that there might be an endogenous substrate for this widely expressed methyltransferase enzyme. The only information presently available supporting this possibility was the report that a methyl acceptor substrate for TPMT was present in the plasma of patients with chronic renal failure (Pazmino et al., 1980). Whether that substrate was endogenous or exogenous was unclear. If there is an endogenous substrate—which remains to be determined—it would be of interest to determine the possible consequences of its decreased methylation in subjects with genetically low levels of TPMT activity.

Other open questions relate to possible additional molecular genetic mechanisms that might participate in the regulation of levels of TPMT activity beyond those that are presently known. For example, in one large population study, 16% of samples with RBC TPMT activity levels that were phenotypically within the heterozygous range had no SNPs within TPMT exons or at exon-intron splice junctions (Otterness et al., 1997). Therefore, the question of additional 5'-flanking region variants beyond the polymorphic VNTR that is presently known, or even of functionally significant intron-based gene sequence variants, remains unanswered. Finally, the issue of the mechanism responsible for the induction of TPMT activity during long-term therapy of patients with leukemia (Lennard et al., 1990) remains to be addressed. Both of these latter issues might have clinical implications. For example, if some patients can induce this pathway for thiopurine metabolism while others cannot, the potential therapeutic implications are obvious. In summary, it is premature to conclude that our present level of understanding of TPMT molecular biology is complete.

Conclusions

The TPMT genetic polymorphism represents a striking example of the potential impact of pharmacogenetics on medicine. The use of thiopurine drugs, a class of drugs that resulted from one of the earliest applications of rational drug design to drug development (Elion, 1967), continues to increase. These agents are widely used in the treatment of neoplastic disease, autoimmune disease, and organ transplant recipients (Paterson and Tidd, 1975; Lennard, 1992). However, because thiopurines—like many other antineoplastic agents—have a relatively narrow therapeutic index, individualization of therapy with these agents is critical. Furthermore, thiopurines, like most xenobiotics, have multiple pathways of biotransformation (Fig. 2). However, it has become clear over the past two decades that one important factor responsible for variation in the metabolism, toxicity, and therapeutic efficacy of these drugs is the genetically polymorphic S-methylation pathway catalyzed by TPMT (Weinshilboum and Sladek, 1980). The fact that systematic study of TPMT in humans was initiated with the expressed goal of exploring the possibility of pharmacogenetic variation but that it has required approximately two decades for this field of research to evolve to its present state—both in terms of clinical implications and molecular understanding of the polymorphism—is a useful lesson with regard to the time required for the translation of pharmacogenetic information into clinical practice. However, because of increasing recognition of pharmacogenetic influence on enzymes that, like TPMT, catalyze pathways of biotransformation, as well as
other proteins that might influence drug response, the possibility that drug therapy might be individualized in part on the basis of pharmacogenetic information has increased rapidly. A major challenge for the future will be finding ways to accelerate both the discovery process and the translation of pharmacogenetic information into meaningful, cost-effective clinical reality.

Acknowledgment. I thank Luanne Wussow for assistance with the preparation of this manuscript.

References


