Oncogene (2006) 25, 1629–1638

www.nature.com/onc

REVIEW

Thiopurine *S*-methyltransferase pharmacogenetics: insights, challenges and future directions

L Wang and R Weinshilboum

Division of Clinical Pharmacology, Department of Molecular Pharmacology and Experimental Therapeutics, Mayo Clinic College of Medicine, Rochester, MN, USA

The thiopurine *S*-methyltransferase (TPMT) genetic polymorphism is one of the most 'mature' examples in pharmacogenetics. That is true because of its importance clinically for the individualization of thiopurine drug therapy and also because TPMT has provided novel insights into molecular mechanisms responsible for the functional effects of common genetic polymorphisms. This review will summarize the development of our understanding of the role of inheritance in the regulation of TPMT as well as the clinical implications of that genetic regulation. It will also summarize recent studies in which TPMT pharmacogenetics has enhanced our understanding of molecular mechanisms by which common polymorphisms influence or alter function. TPMT pharmacogenetics highlights the potential clinical importance of the translation of pharmacogenetics from bench to bedside, the potential for basic pharmacogenetic research to provide insight into mechanisms by which genetic polymorphisms can alter function, and the challenges associated with the achievement of both of those goals.
Oncogene (2006) **25,** 1629–1638. doi:10.1038/sj.onc.1209372

Keywords: TPMT; 6-mercaptopurine; azathioprine; pharmacogenomics; proteasome degradation; aggresome

Introduction

Pharmacogenetics is the study of the role of inheritance in individual variation in drug response – with adverse drug reactions at one end of the spectrum and inadequate therapeutic response at the other. Thiopurine *S*-methyltransferase (TPMT) represents one of the most striking examples of the potential of pharmacogenetics to contribute to individualized drug therapy, in this case therapy with thiopurine drugs such as 6-mercaptopurine (6-MP) and azathioprine (Lennard *et al*., 1987, 1989, 1990). The concept that inheritance can play an important role in variation in drug response originally arose from clinical observations of striking differences among patients in response to 'standard' doses of a drug. Attempts to understand that variation led, sequentially, from clinical phenotypes to identification of the protein(s) responsible and – eventually – to the gene encoding that protein, as well as functional polymorphisms in that gene (Weinshilboum, 2003). During the formative years of pharmacogenetics, application of this 'phenotype-to-genotype' research strategy most often involved drug-metabolizing enzymes like TPMT – enzymes capable of influencing the concentration of drug reaching its target, that is, pharmacokinetics, but pharmacogenetics is increasingly shifting its focus to the drug target itself, that is, to pharmacodynamics (Weinshilboum and Wang, 2004b).

TPMT represents one of the most 'mature' examples of pharmacogenetic research, an example that has been studied for a quarter of a century and is being increasingly applied at the bedside. For that reason, TPMT was highlighted by the United States Food and Drug Administration (FDA) as one of the two 'valid biomarkers' for pharmacogenetics and pharmacogenomics in the 2003 FDA 'Draft Guidance for Pharmacogenomic Data Submission' (US Department of Health and Human Services Food and Drug Administration, 2003). In the following paragraphs, we will briefly review the process by which the TPMT genetic polymorphism was discovered and characterized, evidence in support of the clinical significance of that polymorphism and recent information with regard to mechanisms by which polymorphisms in *TPMT* alter function – with a focus throughout on both future research directions and remaining challenges. The hope is that the quarter century of experience with this 'posterchild' for pharmacogenetics contains lessons that might accelerate future pharmacogenetic research. TPMT also serves to highlight how much progress has been made, as well as how much remains to be learned – even with regard to this very well-studied example of pharmacogenetic science.

TPMT genetic polymorphism: discovery and clinical significance

TPMT pharmacogenetics: background and genomic understanding

TPMT is a cytosolic drug-metabolizing enzyme that catalyses the S-methylation of cytotoxic and immuno-

Correspondence: Dr L Wang, Department of Molecular Pharmacology and Experimental Therapeutics, Mayo Clinic College of Medicine, Rochester, MN 55905 USA.
E-mail: wang.liewei@mayo.edu

suppressant thiopurine drugs such as 6-mercaptopurine (6-MP) and azathiopurine, a prodrug that is converted to 6-MP *in vivo* (Remy, 1963; Woodson and Weinshilboum, 1983; Lennard, 1992). Thiopurines are widely used to treat acute lymphoblastic leukemia (ALL) of childhood, inflammatory bowel disease, autoimmune diseases and organ transplant recipients (Paterson and Tidd, 1975; Lennard, 1992). These drugs, like many cytotoxic agents, have a relatively narrow therapeutic index, with the potential for life-threatening drug-induced toxicity, primarily myelosuppression (Lennard *et al*., 1989; Evans *et al*., 1991; Lennard, 1992). Large inherited variations in human tissue TPMT enzyme activity – ranging from high to virtually undetectable levels of activity – were first described over two decades ago (Weinshilboum and Sladek, 1980) (Figure 1). The first human tissue studied was the red blood cell (RBC), and Figure 1 shows the trimodal frequency distribution histogram of RBC TPMT activity in 298 control subjects as reported in that original study, a frequency distribution that was shown by family studies to be due to inheritance (Weinshilboum and Sladek, 1980). Levels of TPMT activity in the RBC reflect relative levels of that enzyme in the kidney, liver and lymphocyte (Van Loon and Weinshilboum, 1982; Woodson *et al*., 1982; Szumlanski *et al*., 1992; Coulthard *et al*., 1998). After the human TPMT cDNA and gene had been cloned and characterized (Honchel *et al*., 1993; Szumlanski *et al*., 1996; Tai *et al*., 1996), these phenotypic variations were shown to result primarily from variation in the sequence of the TPMT gene itself. A total of 21 *TPMT* genetic polymorphisms have been identified which are, or may be associated with decreased levels of TPMT enzyme activity and/or thiopurine drug-induced toxicity (Salavaggione *et al*., 2005). Eighteen of those polymorphisms (**2*, **3A*, **3B*, **3C*, **5*–**14* and **16*–**19*) involve nonsynonymous coding single-nucleotide polymorphisms (cSNPs) (Krynetski *et al*., 1995; Szumlanski *et al*., 1996; Otterness *et al*., 1997; Lindqvist *et al*., 2004; Schaeffeler *et al*., 2004; Hamdan-Khalil *et al*., 2005), that is, alterations in single DNA nucleotides that alter the encoded amino acid. The **14* SNP disrupts the translation initiation codon (Lindqvist *et al*., 2004) and prevents translation of the enzyme protein (Salavaggione *et al*., 2005), while **4* and **15* involve alterations in canonical mRNA splice site sequences (Otterness *et al*., 1998; Lindqvist *et al*., 2004) and **3D* contains a premature stop codon (Otterness *et al*., 1997). There is also a polymorphic GC-rich variable number of tandem repeats (VNTR) present in the 5′-flanking region of *TPMT* that can vary from 3 to 9 repeat elements (Spire-Vayron de la Moureyre *et al*., 1998, 1999; Yan *et al*., 2000). This VNTR has been reported to modulate TPMT activity as a result of altered transcription.

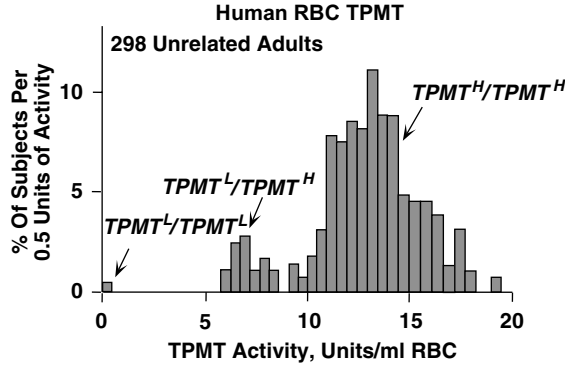

Figure 1 The TPMT genetic polymorphism. The figure shows a frequency distribution histogram of level of RBC TPMT enzyme activity in blood samples from 298 randomly selected Caucasian blood donors. Presumed genotypes for the TPMT genetic polymorphism are also indicated. These designations for high ($TPMT^{H}$) and low ($TPMT^{L}$) activity alleles were used before the molecular basis of the polymorphism was determined. The major alleles responsible were subsequently determined to be *TPMT*3A* and *TPMT*3C* (see Figure 2 and text). Reproduced with permission of the University of Chicago Press (Weinshilboum and Sladek, 1980).

*TPMT *3A*, the most common variant allele in Caucasians (frequency approximately 5%), has two nonsynonymous cSNPs that result in Ala154Thr and Tyr240Cys alterations in encoded amino acids. *TPMT *3C*, the most common functionally significant variant allele in East Asia (frequency approximately 2%), includes only the codon 240 SNP, and the rare *TPMT *3B* allele has only the codon 154 SNP (Figure 2) (Szumlanski *et al*., 1996). The presence of *TPMT*3A* and **3B* result in a virtual lack of TPMT enzyme activity and protein in the tissues of subjects who carry these alleles and, as a result, patients homozygous for these alleles can suffer severe, life-threatening myelosuppression when treated with standard doses of thiopurines, that is, they are 'overdosed' on standard doses (Lennard *et al*., 1989, 1990; Evans *et al*., 1991). The reason why altering only two amino acids out of 245 results in such a striking phenotype will be addressed subsequently. *TPMT*3C* and the first variant allele identified, *TPMT*2*, do not result in such dramatic decreases in levels of enzyme protein as do **3A* and **3B*, but they are also associated with significant decreases in quantity of TPMT protein (Tai *et al*., 1997, 1999; Wang *et al*., 2003; Salavaggione *et al*., 2005). On the basis of population studies, **3A* and **3C* are the predominant variant alleles, with *2 contributing to a lesser extent. These three alleles account for over 95% of cases of inherited TPMT deficiency in Caucasian subjects (McLeod *et al*., 2000). Inherited decrease in TPMT results in potentially life-threatening clinical consequences because of the way in which thiopurines are thought to exert their cytotoxic therapeutic effects. The clinical implications of polymorphisms in *TPMT* will be addressed in subsequent paragraphs.

TPMT pharmacogenetics: thiopurine molecular pharmacology

The immunosuppressant agent azathioprine is a prodrug that is converted to 6-MP *in vivo*, and 6-MP itself is also a prodrug, undergoing a series of enzymatic reactions to form 6-thioguanine nucleotides (6-TGNs), active metabolites that can be incorporated into DNA (Paterson and Tidd, 1975; Lennard *et al*., 1983; Lennard, 1992)

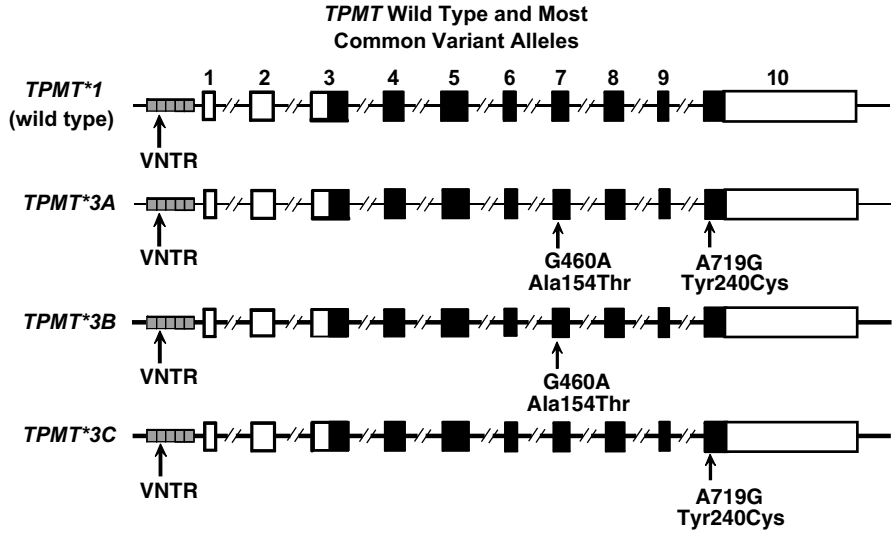

Figure 2 TPMT alleles. *TPMT*1* is the most common allele ('wild type'), while *TPMT*3A* is the most common variant allele in Caucasian subjects and *TPMT*3C* is the most common variant allele in East Asian subjects. Black rectangles represent the open reading frame (ORF) while open rectangles represent 5′- and 3′-untranslated region (UTR) sequence. 'VNTR' represents a GC-rich variable number of tandem repeats that is located in the 5′-flanking region of the gene.

(Figure 3). Two major metabolic reactions inactivate thiopurine drugs *in vivo*, oxidation catalysed by xanthine oxidase and methylation catalysed by TPMT (Weinshilboum, 2001; Krynetski and Evans, 2003; Weinshilboum, 2003). Although the bulk of thiopurines are metabolized by xanthine oxidase (XO), variation in the methylation step is much greater than for XO-catalysed oxidation (Guerciolini *et al*., 1989; Krynetski and Evans, 2003) and, in hematopoietic tissue, the methylation pathway is critical because xanthine oxidase is not expressed in hematopoietic tissue (Lennard *et al*., 1987). A series of clinical studies have shown that level of RBC TPMT activity is inversely correlated with the level of RBC 6-TGN and, as a result, inherited decrease in the methylation step makes more drug available for the multistep pathway that leads to the formation of active metabolites, 6-TGNs (see Figure 3) – resulting in drug-induced myelosuppression (Lennard *et al*., 1987, 1989, 1990; Evans *et al*., 1991). Therefore, patients homozygous for alleles that result in low TPMT activity have elevated 6-TGNs when treated with standard doses of thiopurines and are at greatly increased risk for life-threatening, drug-induced myelosuppression. To avoid toxicity, these patients must be treated with 1/10th to 1/15th of standard doses of thiopurine and, even then, they must be monitored carefully (Weinshilboum, 2001; Krynetski and Evans, 2003; Weinshilboum, 2003). It should be emphasized that there are many possible causes for myelosuppression other than an inherited decrease in TPMT in patients with diseases like ALL – patients who are treated with several cytotoxic drugs. However, there can no longer be any doubt that exposure of a patient with inherited lack of TPMT to standard doses of thiopurines represents a major risk for the rapid development of bone marrow suppression. Some authors have complained that the TPMT polymorphism fails to explain all cases of myelosuppression in these patients (Nebert *et al*., 2003). That criticism is puzzling since 'certainty' is rare in clinical medicine and very few diagnostic tests explain all cases of a particular adverse drug reaction. The fact that TPMT pharmacogenetic testing would allow us to understand, anticipate and avoid this potentially fatal drug reaction in a subset of patients clearly represents a significant clinical advance. Inherited deficiency of TPMT has also been associated with increased risk for radiation-induced

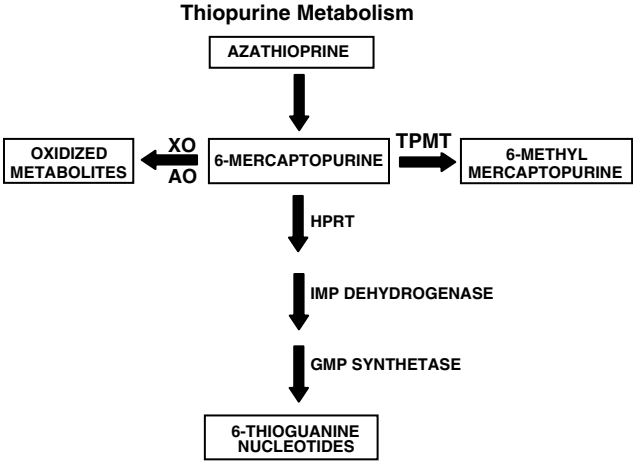

Figure 3 Thiopurine metabolism. The figure shows a simplified schematic representation of the metabolism of the thiopurine drugs azathioprine – which is converted to 6-mercaptopurine (6-MP) *in vivo* – as well as the biotransformation of 6-MP by TPMT, xanthine oxidase (XO) and aldehyde oxidase (AO). The 'metabolic activation' of 6-MP to form 6-thioguanine nucleotides occurs as a result of a series of reactions catalysed by hypoxanthine guanine phosphoriboxyltransferase (HPRT), IMP dehydrogenase and GMP synthetase.

brain tumors in patients treated with thiopurines and with radiation therapy (Relling *et al*., 1999). Conversely, there is evidence that patients with high levels of TPMT may not respond as well to thiopurine drug therapy (Lennard *et al*., 1990) and may require treatment with higher than standard doses, as well as very recent evidence that *TPMT* gene duplication can occur in leukemia cells – with possible implications for drug response (Cheng *et al*., 2005). Finally, it should be emphasized that this brief clinical overview and the schematic metabolic pathway shown in Figure 3 have both been highly simplified. Many details and many observations requiring additional study could not be included. For example, *S*-methylmercaptopurine is not merely an inactive metabolite since it is capable of inhibiting purine biosynthesis (Tay *et al*., 1969; Hill and Bennett Jr, 1980; Dervieux *et al*., 2001). Furthermore, there is evidence that patients who have very high levels of the *S*-methyl metabolite after treatment with 6-MP are also at risk for an adverse response to thiopurine therapy. However, in this case, the adverse response is hepatotoxicity (Dubinsky *et al*., 2002). Obviously, much remains to be learned about thiopurine metabolism and effect in patients. This is merely one of the many challenges that remain if we are to translate even a well-developed pharmacogenetic model such as that represented by TPMT into the clinic.

TPMT pharmacogenetics: clinical testing

A series of important lessons might be learned from the 25-year odyssey of TPMT pharmacogenetics, a journey from a laboratory-based hypothesis to what today has become a widely applied, commercially available diagnostic test. First, the existence of ongoing, clinical trials for ALL made it possible to 'test' that laboratory-based hypothesis quickly and rigorously – in the case of the initial studies, in collaboration with Drs Lynne Lennard and John Lilleyman through their involvement with the United Kingdom Acute Lymphoblastic Leukemia (UKALL) trials (Lennard *et al*., 1987, 1990). The second critical factor was the availability of an intermediate phenotype', the measurement of RBC 6-TGNs, originated and validated by Dr Lennard (Lennard *et al*., 1983) – a critical link between thiopurine drug exposure and the clinical phenotype of myelosuppression. Therefore, if the TPMT pharmacogenetic model has anything to offer to the world of translational research, it might include the wider use of clinical trials to test pharmacogenetic-pharmacogenomic hypotheses and the development and application of additional intermediate phenotypes to sharpen and narrow clinically defined end points (e.g., 'myelosuppression'). However, in spite of the fact that the TPMT genetic polymorphism represents one of the best-validated examples of the clinical importance of pharmacogenetics, many issues still remain to be resolved. Specifically, the question of whether assays for TPMT genotype, phenotype (i.e., RBC TPMT activity and/or RBC 6-TGN concentrations) or both should be performed remains open – as do questions of whether clinical guidelines appropriate for one disease, for example, ALL, are applicable to other pathologic states such as inflammatory bowel disease.

Development of a sensitive radiochemical assay for the measurement of TPMT activity (Weinshilboum *et al*., 1978) – presumably the phenotype of ultimate interest in the clinic – made it possible to perform the original population and family studies of TPMT activity, studies that resulted in the discovery of the TPMT genetic polymorphism (Weinshilboum and Sladek, 1980). The measurement of RBC TPMT activity can be used as a guide for the individualization of thiopurine therapy (see Figure 1), and this test has been applied clinically for that purpose since at least the early 1990s (Laboratory Medicine Bulletin Mayo Foundation, 1991). Furthermore, as pointed out previously, there is a strong correlation between relative levels of TPMT activity in the RBC and in other human tissues (Woodson *et al*., 1982; Szumlanski *et al*., 1992). Following the cloning of the *TPMT* gene and characterization of its most important polymorphisms (Szumlanski *et al*., 1996; Tai *et al*., 1996), determination of *TPMT* genotype was also shown to be useful for individualizing thiopurine drug therapy. However, both genotyping and phenotyping have their limitations. Obviously, if a patient has received transfusions, phenotyping by measuring RBC TPMT activity can be misleading. On the other hand, we still do not understand all DNA sequence variation that influences TPMT enzyme activity, so genotyping – at least at present – is also not foolproof. However, because of the life-threatening nature of thiopurine drug-related toxicity, prospective identification of patients with decreased TPMT activity prior to the initiation of therapy has increasingly been accepted clinically – most recently primarily by genotyping. As pointed out earlier, in addition to the assay of level of RBC TPMT enzyme activity, measurement of RBC 6-TGN represents an additional 'phenotype' that can be used to guide therapy. RBC 6-TGN concentrations are inversely related to levels of TPMT enzyme activity and are directly related to drug-induced toxicity (Lennard *et al*., 1987, 1989). Therefore, serial RBC 6-TGN assays have also been used to monitor thiopurine efficacy and/or toxicity (Gearry and Barclay, 2005). As a result, debate continues with regard to whether or when to assay RBC TPMT enzyme activity, RBC 6-TGN concentrations, *TPMT* genotype – or some combination of all three – in the clinic. That question remains to be resolved, although genotyping or phenotyping to prevent exposure of patients with very low TPMT to standard doses of thiopurine drugs is finding increasing clinical acceptance (Gearry and Barclay, 2005).

As might be anticipated, many techniques have been used to genotype *TPMT*, beginning with RFLP assays and later extending to allele-specific amplification, direct sequencing, SSCP, DHPLC and – more recently – use of a variety of high throughput platforms (Yates *et al*., 1997; Spire-Vayron de la Moureyre *et al*., 1998; Schaeffeler *et al*., 2003, 2004). However, beyond all of the usual issues associated with genotyping, TPMT pharmacogenetics serves to highlight – in a very

practical sense – a limitation common to all of these genotyping methods, their inability to determine haplotype directly. Specifically, because the most common clinically relevant variant allele in Caucasians, *TPMT*3A*, has two functionally significant SNPs separated by 8310 bps, it is important to differentiate the relatively common *1/*3A genotype (approximately 10% of all samples in Caucasians, Figure 1) – that is associated with intermediate activity – from the rare *3B/*3C genotype that is associated with very low activity and high risk for severe drug toxicity (see Figure 2). Stated simply, it would be important clinically to determine haplotype, that is, does a specific patient have one WT allele and another allele with two SNPs (intermediate TPMT activity) or two variant alleles – each bearing a single 'disabling' SNP – with resultant extremely low activity. Although the latter situation is very rare, the consequences of treating such a patient with standard doses of thiopurines could be devastating. In this case, the example provided by TPMT serves to place the focus on a limitation of virtually all methods for genotyping that are used today. To address this problem for TPMT, molecular haplotyping methods have been developed that can be used to distinguish the *1/*3A and *3B/*3C genotypes (McLeod *et al*., 2000; Krynetski and Evans, 2003). Finally, even though more than 95% of functionally important variant alleles are accounted for by **2*, **3A, *3B* and **3C*, rare polymorphisms obviously would not be detected by genotyping that is directed only at known variants.

The TPMT genetic polymorphism also serves to illustrate, in a striking fashion, the fact that allele frequencies and types often vary greatly among different ethnic groups. Population studies conducted among European, Caucasian, East and West African and East Asian populations have demonstrated significant ethnic differences in TPMT pharmacogenetics (Otterness *et al*., 1997; Collie-Duguid *et al*., 1998; Hon *et al*., 1999). *TPMT *3A* is the most common variant allele in Caucasians, with a frequency of approximately 5%, and – as a result – this allele accounts for the majority of variant alleles in white populations in the US and Northern Europe. However, *TPMT*3C* accounts for approximately 50% of variant alleles in African-Americans (Hon *et al*., 1999) and the **3C* allele is the major variant allele in East Asian populations – populations that generally lack *TPMT*3A* (Collie-Duguid *et al*., 1998). This striking difference in allele frequencies and types among different populations could have significant impact on the clinical application of individualized drug therapy. The issue of ethnic variation in drug response has also been highlighted by the recent approval of the drug BiDil for the treatment of severe congestive heart failure in a single ethnic group, African-Americans – a group who displayed a 42% reduction in mortality, while no significant benefit could be demonstrated for Caucasian-American patients (Branca, 2005). TPMT has also served as a focus for basic pharmacogenetic research, just as it has served as a model system for translational pharmacogenetic research.

TPMT: functional genomics and molecular mechanisms

In addition to its clinical importance, and in part as a result of its clinical importance, TPMT has served as an important 'model system' for use in studies of mechanisms responsible for the functional effects of genetic polymorphisms. *TPMT* includes a series of polymorphisms within different regions of the gene, including exons, exon–intron splice junctions and the 5′-flanking region (5′-FR). Each of these alterations in DNA sequence has been shown to have functional effects on enzyme activity. However, most studies of *TPMT* have focused on nonsynonymous cSNPs – polymorphism within the open reading frame (ORF) that alter the encoded amino acids – and those studies have opened new areas of basic pharmacogenetic research. In the following paragraphs, we will discuss *TPMT* nonsynonymous cSNPs and our evolving understanding of mechanisms responsible for their effects, followed by a brief discussion of other functional polymorphisms in this important gene.

TPMT pharmacogenetics: gene sequence variation that alters the encoded amino acid

Nonsynonymous cSNPs are easily the most common gene sequence variation associated with functional effects on TPMT. Thirteen of those nonsynonymous cSNPs were recently studied functionally (Salavaggione *et al*., 2005). Specifically, alleles including those polymorphisms were transiently expressed in COS-1 cells and enzyme activity and protein quantity were determined. The results demonstrated that 11 out of the 13 of these naturally occurring inherited alterations in amino-acid sequence had functional implications after expression in a mammalian cell (Salavaggione *et al*., 2005). Although there were changes in substrate kinetics, the functional effects resulted primarily from alterations in level of enzyme protein. Specifically, Figure 4 show that there was a highly significant correlation between level of enzyme activity and quantity of immunoreactive protein in these variant allozymes – ranging from high values for the WT sequence to virtually no protein for TPMT*3A (Salavaggione *et al*., 2005). This phenomenon has also been observed for many other common genetic polymorphisms that alter only a single amino acid (Freimuth *et al*., 2001; Thomae *et al*., 2002; Adjei *et al*., 2003; Thomae *et al*., 2003; Weinshilboum and Wang, 2004a; Ji *et al*., 2005). Of the allozymes studied by Salavaggione *et al*. (2005), TPMT*3A, *3B, *3C and *2 displayed the most striking effects. This phenomenon was first observed a decade ago when TPMT*3A, *3B and *3C were initially described (Szumlanski *et al*., 1996). In the course of those early experiments, these three allozymes, as well as WT TPMT, were expressed in COS-1 cells, and Western blot analysis was performed. Genotype–phenotype correlation studies performed with human liver biopsy samples also showed that the presence of the **3A* allele was associated with decreased levels of TPMT protein in these tissue samples – confirming the observations in transfected cells (Szumlanski *et al*., 1996).

The association of naturally occurring nonsynonymous cSNPs with altered levels of protein appears to be a common explanation for the functional effects of this type of genetic polymorphism. That observation raises the question of exactly how changing only one or two amino acids could alter protein quantity. There are several possible explanations, including decreased messenger ribonucleic acid (mRNA) stability, decreased rate of protein synthesis or accelerated protein degradation. In most cases that have been studied in detail, including TPMT, accelerated degradation has been the major mechanism (Tai *et al*., 1997, 1999; Wang *et al*., 2003). Pulse chase experiments performed with cultured mammalian cells, as well as experiments performed with the rabbit reticulocyte lysate, an experimental system that has been widely used to study protein degradation, demonstrated that the common TPMT*3A variant allozyme was degraded much more rapidly than was the WT enzyme (Tai *et al*., 1997, 1999; Wang *et al*., 2003) (Figure 5). That process involved ubiquitination of the variant allozyme, followed by proteasome-mediated degradation (Tai *et al*., 1997, 1999).

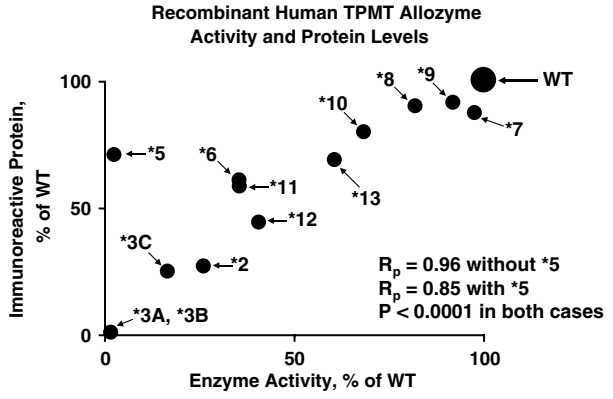

Figure 4 Correlation of recombinant human TPMT allozyme enzyme activity and level of immunoreactive protein after expression in COS-1 cells. WT is the 'wild type' allozyme. The other numbers refer to variant allozyme sequences. Note that the common *3A variant displays neither enzyme activity nor protein. Reproduced with permission of Lippincott Williams and Wilkins (Salavaggione *et al*., 2005).

The next question involved the process by which the cell 'recognizes' an alteration in only one or two amino acids and targets a variant allozyme for proteasome-mediated degradation. In the case of TPMT, there is evidence that molecular chaperones participate in that process. Specifically, heat shock protein 90 (hsp90), hsp70 and heat shock organizing protein (hop) were much more highly associated with the *3A variant allozyme than with the WT protein in the rabbit reticulocyte lysate (Wang *et al*., 2003). In addition, that association had functional consequences for ubiquitination and degradation, at least after treatment of WT TPMT with the hsp90-specific inhibitor geldanamycin (Wang *et al*., 2003). Complexes of proteins involving molecular chaperones and their client proteins participate in a cellular 'protein quality control' system that can either result in proper folding of the client proteins or the targeting of misfolded proteins for proteasome-mediated degradation (Wickner *et al*., 1999; Hohfeld *et al*., 2001; Neckers, 2002; Meusser *et al*., 2005). These misfolded proteins may also aggregate, and the cell has special processes to deal with aggregated protein, processes that can result in aggresome formation. Aggresomes are pericentriolar cytoplasmic structures in which aggregated, polyubiquinated misfolded proteins can be sequestered (Johnston *et al*., 1998; Kopita, 2000; Garcia-Mata *et al*., 2002). Aggresome formation is microtubule dependent, with the involvement of histone deacytalase 6 (HDAC6) and motor proteins such as dynein (Johnston *et al*., 2002; Kawaguchi *et al*., 2003). The fact that TPMT*3A is targeted for accelerated proteasome-mediated degradation, with the involvement of molecular chaperones, raised the possibility that the two common polymorphisms in TPMT*3A might result in misfolding. If that were to occur, TPMT*3A might also aggregate.

Recent studies have demonstrated that the TPMT*3A polymorphisms result in misfolding and protein aggregation – with aggresome formation (Wang *et al*., 2005). Specifically, after transfection into COS-1 cells in the presence of the proteasome inhibitor, MG132,

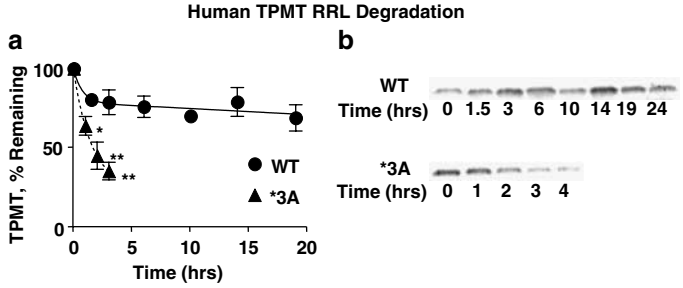

Figure 5 TPMT degradation. ^{35}S-Methionine-labeled WT and *3A TPMT allozymes were incubated in an 'untreated' rabbit reticulocyte lysate (RRL) that included an ATP generating system. Aliquots were removed at the times indicated and subjected to SDS–PAGE, followed by autoradiography. The density of the TPMT band at each time point was expressed as a percentage of the value at zero time. (**a**) Each time point is shown as the mean ± s.e.m. for four experiments. *$P<0.05$ and **$P<0.01$ vs WT at the same incubation time. (**b**) Representative SDS–PAGE for WT and *3A TPMT at different incubation times. Reproduced with permission of Lippincott Williams and Wilkins (Wang *et al*., 2003).

TPMT*3A, but not the WT protein, formed aggresomes (Wang *et al*., 2005) (Figure 6). Furthermore, both the microtubule-destabilizing agent, vinblastine, and the HDAC6 inhibitor, scriptaid, disrupted TPMT*3A aggresome formation, indicating that *3A aggresome formation was microtubule dependent. In addition, size exclusion chromatography of bacterially expressed recombinant WT, *3A, *3B and *3C TPMT confirmed that the presence of the two SNPs in *TPMT*3A* resulted in protein aggregation (Wang *et al*., 2005). These observations indicate that common polymorphisms such as those in TPMT can cause protein misfolding, resulting in both accelerated degradation and aggresome formation. Disruption of protein folding and resultant alteration in the dynamic balance that exists in the protein quality control mechanism may be one process by which nonsynonymous cSNPs alter function in the environment of the mammalian cell. Obviously, these observations would not have been made if only bacterial expression had been performed. It should also be emphasized that mutations or common polymorphisms that result in protein misfolding are not unusual. For example, the most common variant (ΔF508) for the cystic fibrosis transmembrane conductance regulator (CFTR) involves the deletion of a single amino acid, resulting in protein misfolding, accelerated degradation and aggregation (Jensen *et al*., 1995; Johnston *et al*., 1998). However, TPMT represents the only example in pharmacogenetics that has been studied mechanistically in this fashion. There is presently little comparable information available with regard to pharmacogenetically important membrane-bound drug-metabolizing enzymes such as the cytochromes P450 (CYPs).

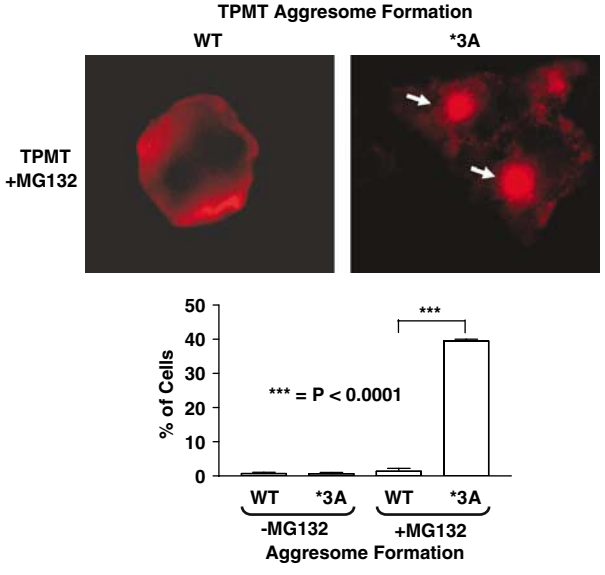

Figure 6 TPMT*3A aggresome formation. (**a**) COS-1 cells were transiently transfected with HA-tagged WT and *3A TPMT constructs in the presence of the proteasome inhibitor MG132 and were subjected to fluorescence microscopy. Arrows point to aggresomes. (**b**) Aggresome formation in a percentage of ~200 cells counted (mean ± s.e.m., $N = 4$). Reproduced with permission from the Proceedings of the National Academy of Science (Wang *et al*., 2005).

The dynamic balance that exists among protein folding, protein degradation and protein aggregation for TPMT is illustrated in Figure 7. Under normal circumstances, WT TPMT can be folded properly, with the involvement of molecular chaperones ('Pathway 1' in Figure 7). However, when the protein contains the two common *3A polymorphisms resulting in Ala154Thr and Tyr240Cys alterations in sequence, the variant allozyme cannot fold properly. This misfolded protein can be polyubiquinated through a series of reactions involving multiple protein–protein interactions, with targeting for proteasome-mediated degradation ('Pathway 2''in Figure 7) (Tai *et al*., 1997, 1999; Wang *et al*., 2003). The misfolded protein can also form microaggregates that can be translocated to aggresomes, with the involvement of microtubules, HDAC6 and the retrograde motor protein, dynein ('Pathway 3' in Figure 7) (Wang *et al*., 2005). All of these processes are in dynamic balance, a balance that can be altered by mutations or polymorphisms – or by an intervention that alters the cellular environment. For example, treatment with the proteasome inhibitor, MG132, can result in blockade of the proteasome-mediated degradation pathway, with a shift to aggresome formation. It is for that reason that virtually all experiments involving the study of aggresomes in cultured cells are conducted in the presence of proteasome inhibitors (Johnston *et al*., 1998; Kopita, 2000).

The observation that TPMT*3A can be targeted for accelerated proteasome-mediated degradation and can also form aggresomes has served as a stimulus to our thinking with regard to mechanisms by which nonsynonymous cSNPs alter function in pharmacogenetics. However, these observations also raise a series of questions with regard to the processes by which cells recognize misfolded proteins, the identity of other

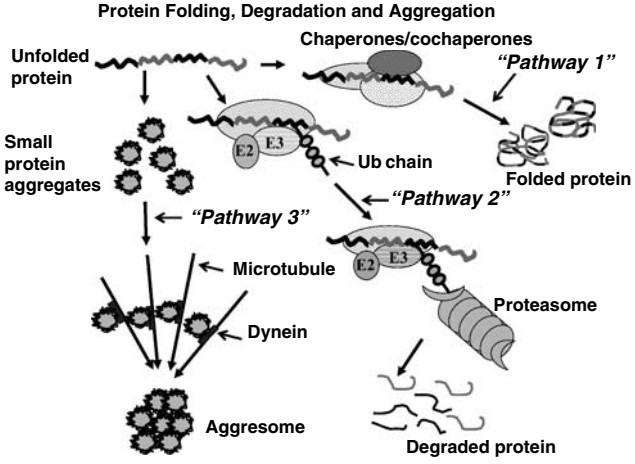

Figure 7 The dynamic balance among protein folding ('Pathway 1') proteasome-medicated degradation ('Pathway 2') and aggresome formation ('Pathway 3'). The figure depicts various 'fates' for a protein, including proper folding, misfolding followed by ubiquitination and proteasome-mediated degradation or aggregation with aggresome formation.

proteins that participate in these processes and how these processes are regulated. All of these mechanisms are of potential practical importance in pharmacogenetics because they could have direct clinical relevance for drug efficacy and/or toxicity.

Finally, there is at least one example of a nonsynonymous cSNP in *TPMT* with an effect on translation. *TPMT*14* has an SNP that results in an A to G transition for the initial nucleotide in the translation initiation codon (Lindqvist *et al*., 2004). This polymorphism results in a Met → Val change in amino acid. That change could potentially activate a downstream translation initiation codon, resulting in a protein that would be 170 amino acids in length instead of 245 (Lindqvist *et al*., 2004), but attempts to translate this altered sequence in the rabbit reticulocyte lysate proved unsuccessful (Salavaggione *et al*., 2005). Finally, even though nonsynonymous cSNPs account for more than 95% of inherited variants in *TPMT* (McLeod *et al*., 2000), polymorphisms located outside of the ORF can also influence TPMT expression. These non-ORF polymorphisms alter both transcription and RNA splicing.

TPMT pharmacogenetics: gene sequence variation that influences transcription

The human TPMT gene does not contain a TATA box or a canonical CAAT sequence (Szumlanski *et al*., 1996). Deletion studies demonstrated that the promoter for *TPMT* was located between (−116) to (−36) bp upstream from the site of transcription initiation (Fessing *et al*., 1998). This region contains a VNTR that consists of GC-rich highly homologous repeats that vary from 17 to 18 bp in length. Spire-Vayron de la Moureyre *et al*. (1999) first described this VNTR in a study of 54 unrelated European individuals. This region contains three to nine repeats, with the most common number in Caucasians being four or five repeats (Spire-Vayron de la Moureyre *et al*., 1999; Yan *et al*., 2000). Reporter gene constructs created with different numbers of repeats showed a modest effect of repeat number on transcription (Spire-Vayron de la Moureyre *et al*., 1999). However, the *in vivo* implications of this VNTR remain controversial. The initial studies showed an inverse relationship between RBC TPMT activity and the sum of repeat element number on both alleles as a surrogate for VNTR genotype (Spire-Vayron de la Moureyre *et al*., 1999). A much larger study performed with over 2000 clinical samples for which RBC TPMT activity had been measured showed a similar, but less clearcut relationship (Yan *et al*., 2000). Specifically, after stratification of the data for the larger study on the basis of RBC TPMT phenotype, there was a significant association of VNTR genotype with level of RBC TPMT activity. The**V4*/**V5* genotype (four repeats/five repeats) was associated with the highest level of RBC TPMT activity, while genotypes that included at least one allele with more than five repeats elements, that is, length > **V5*, were associated with a decreased level of activity (Yan *et al*., 2000). On balance, it appears that the VNTR in the promoter for *TPMT* has only a modest effect on level of enzyme activity. As described subsequently, alterations in splicing – although rare – can also influence TPMT phenotype.

TPMT pharmacogenetics: gene sequence variation that influences mRNA splicing

Sequence variations within introns, especially those that disrupt the canonical splice acceptor sequences, often results in alternative splicing. *TPMT *4*, with a G to A transition at the final splice acceptor nucleotide in intron nine (Otterness *et al*., 1998), and **15* with a G to A transition at the same location in intron seven (Lindqvist *et al*., 2004), are intron-based SNPs that disrupt the intron 9–exon 10 and the intron 7–exon 8, acceptor splice sites. *TPMT *4* results in two different abnormal transcripts, one that results from the activation of a cryptic splice site in intron 9, leading to the inclusion of 330 nucleotides of intron sequence, and another that uses a novel splice site located one nucleotide 3′ downstream from the original splice junction, resulting in transcripts containing a single-nucleotide deletion and a frameshift within exon 10 (Otterness *et al*., 1998). Presence of this allele, which has been studied in an extended kindred, results in very low TPMT activity (Otterness *et al*., 1998). *TPMT*15* results in the loss of 25 amino acids from amino acids 140 to 165, an area of the protein encoded by exon 7 (Lindqvist *et al*., 2004).

TPMT pharmacogenetics: conclusions and future directions

Pharmacogenetics and pharmacogenomics hold out the promise of helping make it possible to achieve the goal of truly individualized drug therapy, and TPMT is one of the most striking examples of the translational potential of pharmacogenetics (Weinshilboum, 2003; Weinshilboum and Wang, 2004b). The TPMT polymorphism has also served to illustrate the importance of pharmacogenetics for broadening our understanding of mechanisms by which common polymorphisms can have functional effects. The rapid advances that have occurred in our ability to link DNA sequence variation to phenotype have provided an opportunity to understand mechanisms responsible for those phenotypic effects – as illustrated by nonsynonymous cSNPs. The data for TPMT, together with many other examples, show that a major mechanism responsible for the functional effects of common inherited alterations in amino-acid sequence is an alteration in protein quantity, most often a decrease (Weinshilboum & Wang, 2004a). Furthermore, those decreases commonly result from accelerated degradation of the variant allozyme, most often through a ubiquitin–proteasome-mediated process. In the case of TPMT, not only is degradation one of the fates for the common TPMT*3A variant allozyme but recent data also indicate that aggresome formation represents another way by which the cell removes abnormal *3A protein (Wang *et al*., 2005). However,

the regulation of these multiple pathways for dealing with misfolded variant allozymes remains unclear. Complexes involving multiple proteins are almost certainly involved in the recognition of the misfolded proteins, but what are the components of those protein complexes? How do they interact? How do these proteins recognize a variant allozyme with only two changes among 245 amino acids? What is the balance between folding, degradation and aggregation, as illustrated in Figure 7, and how is it regulated? Are these processes occurring simultaneously or sequentially, and – if so – in what order? All of these puzzles, and many more, need to be addressed in the future.

TPMT pharmacogenetics has been studied so extensively primarily because of its clinical significance. Knowledge of this drug metabolism polymorphism has already benefited many patients. As a result, TPMT remains one of the few examples in pharmacogenetics that has been successfully translated from the bench to the bedside. In addition, it has served an important function by illustrating the potential value of pharmacogenetics for providing insight into mechanisms by which genetic polymorphisms influence protein function. In the future, pharmacogenetic studies will focus not only on drug metabolism (pharmacokinetics) but also on drug targets (pharmacodynamics) and/or both (Weinshilboum and Wang, 2004b). Finally, even though many questions still remain unanswered, TPMT has clearly served us well as a model system for both translational and basic pharmacogenetics.

Acknowledgements

This study was supported, in part, by National Institutes of Health Grants R01 GM28157, R01 GM35720 and U01 GM61388 (The Pharmacogenetics Research Network). We thank Mrs Luanne Wussow for her assistance with the preparation of this manuscript.

References

Adjei AA, Thomae BA, Prondzinski JL, Eckloff BW, Wieben ED, Weinshilboum RM. (2003). *Br J Pharmacol* **139**: 1373–1382.
Branca MA. (2005). *Nat Rev Drug Discov* **4**: 615–616.
Cheng Q, Yang W, Raimondi SC, Pui CH, Relling MV, Evans WE. (2005). *Nat Genet* **37**: 878–882.
Collie-Duguid ESR, Pritchard SC, Powrie RH, Sludden J, Colier DA, Li T *et al*. (1998). *Pharmacogenetics* **9**: 37–42.
Coulthard SA, Howell C, Robson J, Hall AG. (1998). *Blood* **92**: 2856–2862.
Dervieux T, Blanco JG, Krynetski EY, Vanin EF, Roussel MF, Relling MV. (2001). *Cancer Res* **61**: 5810–5816.
Dubinsky MC, Yang H, Hassard PV, Seidman EG, Kam LY, Abreu MT *et al*. (2002). *Gastroenterology* **122**: 904–915.
Evans WE, Horner M, Chu YQ, Kalwinsky D, Roberts WM. (1991). *J Pediatr* **119**: 985–989.
Fessing MY, Krynetski EY, Zambetti GP, Evans WE. (1998). *Eur J Biochem* **256**: 510–517.
Freimuth RR, Eckloff B, Wieben ED, Weinshilboum RM. (2001). *Pharmacogenetics* **11**: 747–756.
Garcia-Mata R, Gao YS, Sztul E. (2002). *Traffic* **3**: 388–396.
Gearry RB, Barclay ML. (2005). *J Gastroenterol Hepatol* **20**: 1149–1157.
Guerciolini R, Scott MC, Weinshilboum RM. (1989). *FASEB J* **3**: A428.
Hamdan-Khalil R, Gala J-L, Allorge D, Lo-Guidice J-M, Horsmans Y, Houdret N *et al*. (2005). *Biochem Pharmacol* **69**: 525–529.
Hill DL, Bennett Jr LL. (1980). *Biochemistry* **8**: 122–130.
Hohfeld J, Cyr DM, Patterson C. (2001). *EMBO Rep* **2**: 885–890.
Hon YY, Fessing MY, Pui C-H, Relling MV, Krynetski EY, Evans WE. (1999). *Hum Mol Genet* **8**: 371–376.
Honchel R, Aksoy I, Szumlanski C, Wood TC, Otterness DM, Wieben ED *et al*. (1993). *Mol Pharmacol* **43**: 878–887.
Jensen TJ, Loo MA, Pind S, Williams DB, Goldberg AL, Riordan JR. (1995). *Cell* **83**: 129–135.
Ji Y, Salavaggione OE, Wang L, Adjei AA, Eckloff B, Wieben ED *et al*. (2005). *J Neurochem* **95**: 1766–1776.
Johnston JA, Illing ME, Kopito RR. (2002). *Cell Motil Cytoskeleton* **53**: 26–38.
Johnston JA, Ward CL, Kopito RR. (1998). *J Cell Biol* **143**: 1883–1898.
Kawaguchi Y, Kovacs JJ, McLaurin A, Vance JM, Ito A, Yao TP. (2003). *Cell* **115**: 727–738.
Kopita RR. (2000). *Trends Cell Biol* **10**: 524–530.
Krynetski E, Evans WE. (2003). *Oncogene* **22**: 7403–7413.
Krynetski EY, Schuetz JD, Galpin AJ, Pui C-H, Relling MV, Evans WE. (1995). *Proc Natl Acad Sci USA* **92**: 949–953.
Laboratory Medicine Bulletin Mayo Foundation (1991). (December 6 1991). **14**: 1.
Lennard L. (1992). *Eur J Clin Pharmacol* **43**: 329–339.
Lennard L, Lilleyman JS, Van Loon J, Weinshilboum RM. (1990). *Lancet* **336**: 225–229.
Lennard L, Rees CA, Lilleyman JS, Maddocks JL. (1983). *Br J Clin Pharmacol* **16**: 359–363.
Lennard L, Van Loon JA, Lilleyman JS, Weinshilboum RM. (1987). *Clin Pharmacol Ther* **41**: 18–25.
Lennard L, Van Loon JA, Weinshilboum RM. (1989). *Clin Pharmacol Ther* **46**: 149–154.
Lindqvist M, Haglund S, Almer S, Peterson C, Taipalensu J, Hertervig E *et al*. (2004). *Pharmacogenetics* **14**: 261–265.
McLeod HL, Krynetski EY, Relling MV, Evans WE. (2000). *Leukemia* **14**: 567–572.
Meusser B, Hirsch C, Jarosch E, Sommer T. (2005). *Nat Cell Biol* **7**: 766–772.
Nebert DW, Jorge-Nebert L, Vesell ES. (2003). *Am J Pharmacogenom* **3**: 361–370.
Neckers L. (2002). *Trends Mol Med* **8**: S55–S61.
Otterness D, Szumlanski C, Lennard L, Klemetsdal B, Aarbakke J, Park-Hah JO *et al*. (1997). *Clin Pharmacol Ther* **62**: 60–73.
Otterness DM, Szumlanski CL, Wood TC, Weinshilboum RM. (1998). *J Clin Invest* **101**: 1036–1044.
Paterson ARP, Tidd DM. (1975). *6-Thiopurines*. Springer Verlag: New York.
Relling MV, Rubnitz JE, Rivera GK, Boyett JM, Hancock ML, Felix CA *et al*. (1999). *Lancet* **354**: 34–39.
Remy CN. (1963). *J Biol Chem* **238**: 1078–1084.
Salavaggione OE, Wang L, Wiepert M, Yee VC, Weinshilboum RM. (2005). *Pharmacogenetics Genomics* **15**: 801–815.

Schaeffeler E, Fischer C, Brockmeier D, Wernet D, Moerike K, Eichelbaum M *et al*. (2004). *Pharmacogenetics* **14**: 407–417.

Schaeffeler E, Stanulla M, Greil J, Schrappe M, Eichelbaum M, Zanger UM *et al*. (2003). *Leukemia* **17**: 1422–1424.

Spire-Vayron de la Moureyre C, Debuysere H, Fizio F, Sergent E, Bernard C, Sabbagh N *et al*. (1999). *Pharmacogenetics* **9**: 189–198.

Spire-Vayron de la Moureyre C, Debuysere H, Mastain B, Vinner E, Marez D, Lo Guidice JM. (1998). *Br J Pharmacol* **125**: 879–887.

Szumlanski C, Otterness D, Her C, Lee D, Brandriff B, Kelsell D *et al*. (1996). *DNA Cell Biol* **15**: 17–30.

Szumlanski CL, Honchel R, Scott MC, Weinshilboum RM. (1992). *Pharmacogenetics* **2**: 148–159.

Tai H-L, Fessing MY, Bonten EJ, Yanishevsky Y, d'Azzo A, Krynetski EY *et al*. (1999). *Pharmacogenetics* **9**: 641–650.

Tai H-L, Krynetski EY, Schuetz EG, Yanishevski Y, Evans WE. (1997). *Proc Natl Acad Sci USA* **94**: 6444–6449.

Tai H-L, Krynetski EY, Yates CR, Loennechen T, Fessing MY, Krynetskaia NF *et al*. (1996). *Am J Hum Genet* **58**: 694–702.

Tay BS, Lilley RM, Murray AW, Atkinson MR. (1969). *Biochem Pharmacol* **18**: 936–938.

Thomae BA, Eckloff BW, Freimuth RR, Wieben ED, Weinshilboum RM. (2002). *Pharmacogenomics J* **2**: 48–56.

Thomae BA, Rifki OF, Theobald MA, Eckloff BW, Wieben ED, Weinshilboum RM. (2003). *J Neurochem* **87**: 809–819.

US Department of Health and Human Services Food and Drug Administration (2003). Center for Drug Evaluation and Research, Center for Biologics Evaluation and Research & Center for Devices and Radiological Health (November 2003). http://www.FDA.gov.

Van Loon JA, Weinshilboum RM. (1982). *Biochem Genet* **20**: 637–658.

Wang L, Nguyen TV, McLaughlin RW, Sikkink LA, Ramirez-Alvarado M, Weinshilboum RM. (2005). *Proc Natl Acad Sci USA* **102**: 9394–9399.

Wang L, Sullivan W, Toft D, Weinshilboum R. (2003). *Pharmacogenetics* **13**: 555–564.

Weinshilboum R. (2001). *Drug Met Dispos* **29**: 601–605.

Weinshilboum R. (2003). *New Engl J Med* **348**: 529–537.

Weinshilboum R, Wang L. (2004a). *Clin Pharmacol Ther* **75**: 253–258.

Weinshilboum R, Wang L. (2004b). *Nat Rev Drug Disc* **3**: 739–748.

Weinshilboum RM, Raymond FA, Pazmiño PA. (1978). *Clin Chim Acta* **85**: 323–333.

Weinshilboum RM, Sladek SL. (1980). *Am J Human Genet* **32**: 651–662.

Wickner S, Maurizi MR, Gottesman S. (1999). *Science* **286**: 1888–1893.

Woodson LC, Dunnette JH, Weinshilboum RM. (1982). *J Pharmacol Exp Ther* **222**: 174–181.

Woodson LC, Weinshilboum RM. (1983). *Biochem Pharmacol* **32**: 819–826.

Yan L, Zhang S, Eiff B, Szumlanski CL, Powers M, O'Brien JF *et al*. (2000). *Clin Pharmacol Ther* **68**: 210–219.

Yates CR, Krynetski EY, Loennechen T, Fessing MY, Tai HL, Pui CH *et al*. (1997). *Ann Intern Med* **126**: 608–614.