Published in final edited form as:

J Mol Biol. 2008 June 13; 379(4): 803–814.

Four Human Thiopurine S-Methyltransferase Alleles Severely Affect Protein Structure and Dynamics

Karen Rutherford[†] and Valerie Daggett^{†,‡,*}

[†]Department of Biochemistry, University of Washington, Seattle WA 98195-5013

[‡]Department of Bioengineering, University of Washington, Seattle WA 98195-5013

Summary

Thiopurine S-methyltransferase (TPMT) metabolizes cytotoxic thiopurine drugs used in the treatment of leukemia and inflammatory bowel disease. TPMT is a major pharmacogenomic target with 23 alleles identified to date. Several of these alleles cause rapid protein degradation and/or aggregation, making it extremely difficult to study the structural impact of the TPMT polymorphisms experimentally. We, therefore, have performed multiple molecular dynamics simulations of the four most common alleles (*TPMT* 2 (A80P), 3A (A154T/Y240C), 3B (A154T) and 3C (Y240C)) to investigate the molecular mechanism of TPMT inactivation at an atomic level. The A80P polymorphism in *TPMT* *2 disrupts helix α 3 bordering the active site, which breaks several salt-bridge interactions and opens up a large cleft in the protein. The A154T polymorphism is located within the co-substrate binding-site. The larger threonine alters the packing of substrate binding residues (P68, L69, Y166), increasing the solvent exposure of the polymorphic site. This packing rearrangement may account for the complete lack of activity in the A154T mutant. The Y240C polymorphism is located in β -strand 9, distant from the active site. Side-chain contacts between residue 240 and helix α 8 are lost in *TPMT* *3C. Residues 154 and 240 in *TPMT* *3A are connected through a hydrogen-bonding network. The dual polymorphisms result in a flattened, slightly distorted protein structure and an increase in the thiopurine-binding site solvent accessibility. The two variants that undergo the most rapid degradation *in vivo*, *TPMT**2 and *3A, are also the most deformed in the simulations.

Keywords

thiopurine S-methyltransferase; polymorphisms; molecular dynamics; pharmacogenetics; protein structure

Thiopurine drugs (azathiopurine (AZA), 6-mercaptopurine (6-MP), thioguanine (TG)) are cytotoxic agents increasingly used to treat a variety of conditions including acute lymphoblastic leukemia^{1; 2; 3; 4}, inflammatory bowel disease^{5; 6; 7; 8}, Crohn's Disease⁹, and multiple sclerosis¹⁰. While these drugs are effective, they have a relatively narrow therapeutic index and prolonged exposure can lead to myelosuppression and life threatening leucopenia¹¹.

*Corresponding Author: Phone: (206) 685-1510, Email: daggett@u.washington.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Thiopurine *S*-methyltransferase (TPMT, E.C.2.1.1.67) is one of three enzymes responsible for thiopurine metabolism. 23 *TPMT* alleles that result in intermediate to no TPMT activity are present in ~11% of the population^{2; 8; 12; 13; 14; 15; 16; 17; 18; 19; 20}. Because of the extreme inter-individual variation in TPMT activity, genotyping and TPMT activity measurements are routinely performed prior to thiopurine dosing^{1; 21; 22; 23; 24}. Four alleles (*TPMT* *2, *3A, *3B, and *3C) account for ~95% of inherited TPMT deficiency and have been biochemically characterized.

TPMT *3A contains compound A154T and Y240C polymorphisms and is the most common variant allele in Caucasians²⁵. The *3A allele results in a ~400-fold decrease in protein levels and no detectable enzyme activity^{26; 27}. It appears that the dual polymorphisms cause aberrant protein folding and aggresome formation^{27; 28}. Pulse chase experiments show that *TPMT* *3A undergoes ubiquitin-mediated proteasome degradation with a half-life of ~0.67 hrs, approximately 72 times faster than wild-type *TPMT*^{27; 29}. The *TPMT* *3B and *3C alleles contain the individual A154T and Y240C polymorphisms, respectively. The *TPMT* *3B allele results in a 4-fold decrease in protein levels, with a degradation half-life of ~6 hrs, and no catalytic activity²⁶. The Y240C polymorphism destabilizes *TPMT*, decreasing its T_m value from 50 to 39°C^{27; 30}, and its protein and activity levels are decreased compared to wild type²⁹. The co-substrate *S*-adenosylmethionine (SAM) is able to protect the catalytic activity and stability of *TPMT* *3C^{26; 29} while the addition of SAM has no such effect on *TPMT* *3B²⁶. The *TPMT* *2 allele contains an alanine to proline substitution at residue 80 that results in ~100-fold decrease in TPMT activity and very low levels of immunological protein^{12; 26; 29}. *TPMT* *2 has a half-life of 0.35 hours and is proteasomically degraded via a ubiquitin mediated pathway²⁹.

Two crystal structures of wild-type human TPMT bound with co-product *S*-adenosylhomocysteine (SAH) have been solved³⁰ (Figure 1a). Wu *et al.* used the crystal structures to predict the structural impact of the *2, *3A, *3B, and *3C polymorphisms³⁰. They noted that the A80P polymorphism is located in the last turn of helix $\alpha 3$, on the opposite face of the protein from the active site (Figure 1a). Residue 154 is located near the SAM and thiopurine-binding sites, and the Y240C polymorphism is located in strand $\beta 9$ within the central β -sheet core (Figure 1a). Several high-resolution structures of the TPMT variants are needed to completely understand the effects of these polymorphisms at atomic resolution. However, because most TPMT alleles result in either increased degradation or aggregation of the protein, further structural studies are extremely difficult. In the present study, we have performed multiple molecular dynamics (MD) simulations of wild-type protein (*TPMT* *1) and the *TPMT* *2, *3A, *3B, and *3C alleles at 37°C to examine how these polymorphisms affect TPMT structure and dynamics.

Results and Discussion

TPMT is a monomeric, single-domain protein consisting of a 9-stranded core β -sheet sandwiched between two sets of α -helices. Co-product *S*-adenosylhomocysteine (SAH) contacts several residues in $\alpha 1$, $\alpha 5$, $\alpha 6$ and β -strands 1 and 2, and is completely buried in the structure by helix $\alpha 1$ which caps the active site (Figure 1). The position of the SAH S δ atom, and putative position of the transferable CE methyl group, as well as active-site comparisons with structurally similar SAM-dependent methyltransferases (catechol *O*-methyltransferase, histamine *N*-methyltransferase) suggest that the substrate-binding site is situated along the C-terminal β -strands of the protein's core β -sheet. Wu *et al.* (2007) identified a small water channel defined by F40 and P196 that could facilitate docking of a thiopurine substrate (Figure 1). The thiopurine-binding site is bordered by residues in helices $\alpha 1$, $\alpha 6$ and $\alpha 9$ and the loop between $\beta 7$ and $\alpha 8$ and capped by helix $\alpha 1$. Because both substrates appear buried in the TPMT crystal structure, both $\alpha 1$ and the unstructured amino-terminal region must be flexible to allow

substrate binding and SAH:SAM exchange. In general, the overall TPMT fold topology was maintained throughout all of the simulations, and none of the polymorphisms resulted in gross unfolding of the protein at 37°C.

Figure 2 shows the $\text{C}\alpha$ -root-mean square fluctuations ($\text{C}\alpha$ RMSF) about the mean structure for all of the 37°C simulations. The $\text{C}\alpha$ RMSF values from the wild-type (*TPMT* *1) simulations are on the same scale and follow a pattern similar to those of the crystallographic B-factors (Figure 2a). Most structural variation occurred in the loop regions, with the central β -sheet remaining stable. The largely unstructured N-terminus (residues 17–25) had the greatest $\text{C}\alpha$ RMSF values, while large fluctuations were also seen in $\alpha 1$ and $\alpha 5$, $\beta 3$, and the $\beta 3$ - $\beta 4$ loop. These changes resulted mainly from the removal of SAH and solvent molecule 2-[3-(2-hydroxy-1,1-dihydroxymethyl-ethylamino)-propylamino]-2-hydroxymethyl-propane-1,3-diol (B3P) from the TPMT starting structures. B3P is bound in a shallow hydrophilic cleft bordered by $\beta 3$ and $\alpha 5$, where it forms hydrogen bonds with residues in $\beta 1$ (R64), the $\beta 3$ - $\beta 4$ loop (T113, E114), $\beta 5$ (Y131) and $\alpha 5$ (N142) (Figure 1). Residues in $\alpha 5$ contact both SAH (S134, I135) and B3P (N142), preserving its secondary structure despite the presence of a helix breaking proline (P139). The removal of SAH and B3P relaxed the substrate-induced structural constraints on $\alpha 5$, resulting in the reorientation or loss of the helix during the apoprotein simulations while increasing the SAM-binding site solvent-accessible surface area (SASA) by $\sim 140 \text{ \AA}^2$. B3P also forms hydrogen bonds with T113 and E114 in the $\beta 3$ - $\beta 4$ loop, and the loss of these stabilizing contacts allowed the terminal $\beta 3$ and $\beta 4$ strands to pull away from $\alpha 5$, imposing a slight twist in the core β -sheet.

TPMT *2 (A80P) Simulations

Residue 80 is located in $\alpha 3$ (Figure 1, Figure 3a). Substitution of proline for alanine at position 80 breaks the M76 O – A80 N hydrogen bond, disrupting the final turn of $\alpha 3$ (Figures 3a–b). As a result, residue 80 does not form side-chain contacts with residues in $\alpha 3$ (K77, R82) and $\beta 2$ (G83, H84, V86) in *TPMT* *2 (Table 1, Figures 3a–b). These changes result in a large scale ($1.3 \text{ \AA}$) reorientation of $\alpha 3$ coupled with increased fluctuations in $\alpha 2$ and $\alpha 4$ (Table 1, Figure 2a). Pro is known to break or distort helices.

In the wild-type simulations, $\alpha 2$ and $\alpha 4$ are stabilized in part by a series of ionic interactions with $\alpha 3$ (Figure 4a). The repacking of K77 in *TPMT* *2 creates an additional salt bridge (K77-E102) between $\alpha 3$ and $\alpha 4$, stabilizing the new orientations of these helices. This close association with $\alpha 4$ enables P80 to form intermittent interactions with L105 ($\alpha 4$) and S125 ($\beta 4$ - $\beta 5$ loop) not present in the wild-type simulations, and pack more closely against $\beta 5$ (Table 1, Figure 3b). In addition, the loss of $\alpha 3$ helicity in *TPMT* *2 disrupts the D54-R82 salt bridge, increasing the flexibility of $\alpha 2$ and allowing it to move away from $\alpha 3$ (Table 1, Figure 4a).

These local packing changes have a large impact on the active site structure of *TPMT* *2. SAM-binding residues are located on the distal ends of $\alpha 2$ (H41), $\alpha 3$ (L69, A73), $\beta 2$ (E90), and $\beta 5$ (S134, I135). The A80P polymorphism alters the orientation and increases the flexibility of several SAM-binding residues. In addition, structural variations in $\alpha 2$ are translated to $\alpha 1$, which contains three SAM-binding residues (L26, W29, W33) and forms a partial cap over the active site. All of these changes result in the formation of a large cleft in *TPMT* *2, increasing the solvent accessible surface area of the active-site by $\sim 200 \text{ \AA}^2$ (Table 1, Figure 4).

TPMT *2 is rapidly degraded *in vivo*, resulting in very low levels of immunological protein and a 100-fold decrease in enzymatic activity^{12; 26; 29}. The loss of $\alpha 3$ causes a large-scale rearrangement of nearby $\alpha 2$ and $\alpha 1$, destabilizing and distorting the overall protein structure (Figure 4b). This also results in a large increase in active-site solvent exposure that may prohibit the binding of co-substrate SAM thereby significantly decreasing any thiopurine methyltransferase activity.

TPMT *3B (A154T) Simulations

Residue 154 is located in $\alpha 6$ bordering the active site where it forms second-shell contacts with several key SAM-binding residues (L69, G153, A157) (Figure 1). Substitution of threonine for alanine does not have a significant effect on the solvent accessibility of the overall protein or the SAM-binding site (Table 2). The A154T mutation replaces a small hydrophobic residue with a larger hydrophilic residue in the active site of the TPMT protein. Changes in the active site can have drastic effects on catalytic activity, although the destabilizing effect of A154T is less intuitive and here MD helps to explain the effect. Figure 2b shows that there are no significant differences in the mean structures of *TPMT* *1 and *3B in the immediate vicinity of residue 154. However, local structural changes caused by the A154T polymorphism are translated to $\alpha 1$ and $\alpha 9$, which cap the SAM- and thiopurine- binding sites, respectively (Figure 2b).

In the initial structure, A154 forms main-chain hydrogen bonds with D151, A157 and I158 and hydrophobic contacts with residues in $\beta 1$ (P68, L69), $\alpha 6$ (D151, G153, L155, V156, A157) and $\alpha 7$ (Y166) (Figure 3c). These contacts are maintained throughout the simulations even while the polymorphic site loosens up slightly, increasing the solvent accessibility of A154 (Table 2). The hydrophobic interactions of A154 with D151 and Y166 facilitate a hydrogen-bonding network between D151 ($\alpha 6$) - Y166 ($\alpha 7$) - Y180 ($\beta 7$), connecting the SAM-binding site to the C-terminal β -strands, and the putative thiopurine-binding site (Figure 3c).

Substituting the larger, hydrophilic threonine for alanine in $\alpha 6$ both disrupts several interactions mediated by residue 154 (L69, Y166) and increases the solvent exposure of the polymorphic site (Figure 3d, Table 2). The aliphatic CB and CG2 atoms of T154 face the interior of the SAM-binding site maintaining its hydrophobicity, while T154 OG1 forms an additional hydrogen bond with the side-chain carboxyl group of D151. This hydrogen bond disrupts the interactions of Y166 with both D151 and Y180, and allows for the rearrangement of $\alpha 7$ and Y166 in the SAM-binding site (Table 2, Figures 3c–d).

The A154T polymorphism results in reduced levels of immunological protein compared to wild-type TPMT, decreasing its half-life from greater than 24 hours to 5.5 hours^{26; 29}. *TPMT**3B loses activity rapidly at 37°C²⁶, and monomers isolated via size-exclusion chromatography showed no activity²⁷. Interestingly, the addition of co-substrate SAM does not protect this loss of activity²⁶. These studies suggest that the active-site TPMT polymorphism alters both the structure of the protein and its ability to catalyze the methyl-transfer reaction. The T154 induced rearrangement of $\alpha 7$ and increased flexibility of helices $\alpha 1$ and $\alpha 9$ results in a slight distortion of the TPMT structure and may account for the destabilizing effects of the A154T polymorphism. In addition, the disruption of a hydrogen bond network between Y166, D151, and Y180 changes the active-site structure and loss of an interaction between L69 and residue 154 may distort the active-site enough to render the *TPMT**3B protein inactive²⁶.

TPMT *3C (Y240C) Simulations

Residue 240 is located in the centre of $\beta 9$ in the C-terminal portion of the core β -sheet (Figure 1). Y240 forms main-chain hydrogen bonds with both the amide and carbonyl groups of L182 ($\beta 7$) and van der Waals interactions with residues in $\beta 7$ (L182, V184), $\beta 8$ (I214, C216, L217), and $\beta 9$ (K238, L239, L241, L242) to maintain the β -sheet structure. The large tyrosine side-chain also interacts with several residues in $\alpha 8$ (H201, I204, E205) (Figure 3e). The backbone hydrogen bonds and β -sheet interactions are present in *TPMT* *3C. However, several of the side-chain contacts between residue 240 and $\alpha 8$ (H201, E205) are lost upon mutation, causing $\alpha 8$ to pull away from the β -sheet (Figure 3f, Table 2). This loosening up of the C-terminal structure allows slight increases in the C α RMSD of $\alpha 8$ and $\alpha 9$ and in the flexibility of $\beta 8$ - $\alpha 9$

loop (Table 2). However, the Y240C polymorphism has little effect on the fluctuations of the mean *TPMT**3C structure compared to the wild-type protein (Figure 2c) and no effect on total, SAM-site or thiopurine-site solvent accessibility (Table 2).

Experimentally, *TPMT**3C has lower levels of immunological protein and enzymatic activity than the wild-type protein²⁹. In addition, the Y240C polymorphism decreases the T_m of the protein from 50 to 39°C^{27; 30}. Interestingly, the addition of co-substrate SAM both rescues the activity and increases the stability of *TPMT**3C, protecting it from degradation^{26; 29}. This stabilizing effect of SAM is seen in other polymorphic methyltransferases. Catechol *O*-methyltransferase (COMT) has a valine to methionine polymorphism at residue 108, ~16 Å from the SAM binding site. The V108M mutation results in decreased T_{50} ³¹ and T_m values³² compared to the wild-type protein, and 108M COMT is more prone to deactivation through oxidation³³. The addition of SAM protects 108M COMT against both deactivation³⁴ and thermal denaturation.³² MD simulations of 108V and 108M COMT showed that the V108M mutation increases the solvent exposure of the overall protein, distorts the structure of both the SAM and substrate binding sites, and reorients several helices in the vicinity of the polymorphic residue³⁵. These structural effects were minimized in the presence of SAM. In the *TPMT**3C imulations described here, the increase in overall protein solvent accessibility coupled with the rearrangement of helices and loops in the vicinity of the Y240C polymorphism may represent the early stages of *TPMT**3C deformation.

TPMT *3A (A154T/Y240C) Simulations

The *TPMT**3A allele contains dual A154T (α 6) and Y240C (β 9) polymorphisms. The *TPMT* topology was maintained in the *TPMT**3A simulations. However, the protein expanded and became more flattened, with the total solvent accessibility varying dramatically throughout the simulations at physiological temperature (Figure 5, Table 2). The effects of the A154T and Y240C polymorphisms on mean structural fluctuations are additive (Figures 2b–d). Residues 154 and 240 are connected via a hydrogen-bonding network involving side-chain contacts between T154, D151 and Y180, and main-chain interactions between Y180 (β 7)-L242 (β 9) and L182 (β 7)-C240 (β 9) (Figure 5). This connectivity is not present in the simulations of either the wild-type or the single mutant *TPMT**3B or *3C proteins (Figure 5, Table 2). The hydrogen-bonding network facilitates the translation of the Y240C effects on the β -sheet and C-terminal packing to the SAM-binding site, and vice versa.

The Y240C polymorphism in *TPMT**3A resulted in a significant increase in the solvent exposure of the polymorphic site due, primarily, to the lost interactions between residue 240 and H201 and E205 of α 8. The loosened C-terminal packing allowed for the rearrangement of the β 7- α 8 loop and of α 9 connecting β 8 and β 9, both of which border the thiopurine-binding site (Figure 5). As a result of this rearrangement, the thiopurine-binding site solvent exposure increased significantly.

The dual A154T and Y240C polymorphisms appear to have additive destabilizing effects both *in vivo* and *in silico*. *TPMT**3A is rapidly proteasomically degraded with a half life of ~0.3 hours^{26; 27; 28; 29}. This degradation is much too fast to measure *TPMT* activity *in vitro*. In contrast, the *TPMT**3B and 3C proteins have half-lives of 6 and 11 hours, respectively, and there was no loss of immunodetectable protein over the assay period²⁹. The simulations showed that the individual destabilizing effects of the A154T and Y240C polymorphisms were linked in the *TPMT* 3A double mutant through a hydrogen bonding network. This resulted in a much more deformed overall structure than those seen in the individual *TPMT**3B or *3C simulations, suggesting a mechanism for the decreased stability of *TPMT**3A.

Conclusions

We performed molecular dynamics simulations of the *TPMT**2, *3A, *3B, and *3C alleles to study the molecular mechanisms behind their destabilizing effects on TPMT protein structure. Interestingly, while these polymorphisms are distributed throughout the protein structure they all are able to exert an effect on the active site. The A80P (*TPMT**2) mutation disrupts the final turn of $\alpha 3$, altering both local salt bridges and secondary structure packing while increasing the solvent exposure of the SAM-binding site. These structural changes result in the formation of a huge cleft in the protein. Residue 154 is located in the active site and the presence of the A154T (*TPMT**3B) polymorphism reorders several SAM-binding residues, which could ablate thiopurine methyltransferase activity. The overall structures of the wild-type and *TPMT**3C (Y240C) proteins remained similar throughout the simulations, with the exception of helix $\alpha 8$, which was reoriented due to loss of side-chain contacts between its residues and the smaller C240. The compound *TPMT**3A (A154T/Y240C) polymorphisms resulted in deformed and flattened structures, a reordering of catalytic residues within the SAM-binding site and an increase in the solvent exposure of the thiopurine-binding site.

Both *TPMT**2 and *3A are extremely unstable *in vivo* with half-lives of 0.35 and 0.67 hours in COS-1 cells²⁹. Several studies have shown that both *TPMT**2 and *3A are proteasomically degraded via a ubiquitin mediated pathway^{27; 29}. TPMT contains 21 lysine residues of which all 21 are solvent exposed and 12 are found in loops and flexible regions on the protein surface. In our MD simulations, helical rearrangements caused by the A80P (*TPMT**2) polymorphism increased the solvent exposure of K37 and K60 in the $\alpha 1$ - $\alpha 2$ and $\alpha 2$ - $\beta 1$ loops, respectively, relative to the wild-type protein (Figure 6). In addition, the distortion of the *TPMT**3A structure exposed several lysine residues in the protein's amino-terminus (K20, K37) and the loops within the β -sheet (K60, K145, K245) (Figure 6). In general, proteins are more easily degraded by the proteasome when the modified lysine leads into a helix or an irregular surface loop^{36; 37}. Based on their location and the differences seen in simulations of the wild-type and mutant TPMT proteins, it is possible that either K20 or K37 serve as targets for ubiquitin modification in *TPMT**2 and *3A. Mutational analysis of the lysine residues in TPMT coupled with ubiquitination assays are needed to determine which lysine residues are modified by ubiquitin experimentally.

We have described the effects of the four most common TPMT alleles on protein structure and dynamics. However, 23 TPMT alleles affecting either protein activity or stability have been identified to date with the polymorphic residues distributed throughout the protein (Figure 7). Several of these residues are found in either the SAM- (*TPMT**18 (G71R), *TPMT**21 (L69V)) or thiopurine- (*TPMT**7 (H227Q), *TPMT**17 (Q42E)) binding sites. Three alleles appear to cluster in $\beta 8$ (R215H, C216R) and $\beta 9$ (K238E) surrounding the Y240C polymorphism described here. In addition, a Y180F (*TPMT**6) polymorphism exists that potentially affects the hydrogen-bonding network connecting the C-terminal β -strands to the SAM-binding site. Further experimental and computational studies are needed to characterize the structural effects of these TPMT polymorphisms.

Materials and Methods

Protein Preparation

The 1.6 Å crystal structure of human thiopurine *S*-methyltransferase (2BZG.PDB 30, residues 17–245) bound with *S*-adenosylhomocysteine (SAH) and 2-[3-(2-hydroxy-1,1-dihydroxymethyl-ethylamino)-propylamino]-2-hydroxymethyl-propane-1,3-diol (B3P) was the starting structure for all of the simulations. SAH and B3P were removed, and the selenomethionine residues at positions 76, 148 and 170 were replaced with methionine. The *2 (A80P), *3A (A154T, Y240C), *3B (A154T), and *3C (Y240C) TPMT alleles were

generated by replacing each wild-type residue with its respective polymorphic variant, and minimizing the torsional, electrostatic and van der Waals energies of the resultant structures *in vacuo*.³⁸ The root-mean-square deviation of the C α atoms (C α RMSD) between the starting structures of wild-type TPMT and all of the variant proteins was ~0.1 Å.

Molecular Dynamics Simulations

Molecular dynamics (MD) studies of the wild-type and polymorphic TPMT proteins were performed with the *in lucem* molecular mechanics (*ilmm*) simulation package³⁹ using protocols and a potential energy function that have been previously described^{40; 41; 42}. The simulations included all hydrogen atoms and explicit flexible three-centered waters⁴². Proteins were solvated in a periodic rectangular box with walls extending at least 10 Å from all protein atoms. Solvent densities were set to 0.997 and 0.933 g/ml for simulations performed at 25 and 37°C, respectively⁴³. Once the densities were set, the box volume was held fixed and the NVE microcanonical ensemble employed. A 10 Å force-shifted, non-bonded cut off was used and updated every 2 steps⁴⁴. A time step of 2 fs was used in all calculations, and all simulations were run for 21 ns with structures saved every 1 ps for analysis. One 25°C simulation and three independent simulations at 37°C were performed for each system for a total of 420 ns of simulation time.

Analyses

Average values for C α RMSD, solvent-accessible surface area (SASA), and contact distances were calculated using structures from the last 5 ns (5,000 structures) of each simulation. A contact was defined as a C-C atom distance within 5.4 Å, or an other heavy atom distance within 4.6 Å between two non-neighboring residues. SASA was calculated using in-house software implementing the NACCESS algorithm⁴⁴. C α RMSF values plotted in Figure 2 were calculated relative to the average structure over the last 10 ns of each simulation at 37°C. C α RMSF values for the TPMT crystal structure (2BZG.PDB 30) were calculated using the crystallographic B-factors: $C\alpha \text{ RMSF} = (3B/8\pi^2)^{1/2}$.

The data described in Table 1 and Table 2 are based on sets of three independent simulations for each wild-type and variant protein. The errors are the standard deviations in the average values of the ensembles for each property reported. Statistical significance of the data was determined using the Student's *t* test.

Acknowledgements

Financial support for this work was provided by the National Institutes of Health (GM50789 to VD) and the Canadian Institutes of Health Research (DRA MD-75910 to KR). Figures were produced using the USCF Chimera package from the Computer Graphics Laboratory, University of California, San Francisco (supported by NIH P41 RR-01081)⁴⁶. K.R. thanks Drs. William Parson and Richard Weinshilboum for their helpful discussions, and Amanda Jonsson for running the simulations.

References

1. Deeken JF, Figg WD, Bates SE, Sparreboom A. Toward individualized treatment: prediction of anticancer drug disposition and toxicity with pharmacogenetics. *Anticancer Drugs* 2007;18:111–126. [PubMed: 17159598]
2. Schaeffeler E, Stanulla M, Greil J, Schrappe M, Eichelbaum M, Zanger UM, Schwab M. A novel TPMT missense mutation associated with TPMT deficiency in a 5-year-old boy with ALL. *Leukemia* 2003;17:1422–1424. [PubMed: 12835738]
3. Stanulla M, Schaeffeler E, Flohr T, Cario G, Schrauder A, Zimmermann M, Welte K, Ludwig WD, Bartram CR, Zanger UM, Eichelbaum M, Schrappe M, Schwab M. Thiopurine methyltransferase (TPMT) genotype and early treatment response to mercaptopurine in childhood acute lymphoblastic leukemia. *Jama* 293:1485–1489.

4. McLeod HL, Coulthard S, Thomas AE, Pritchard SC, King DJ, Richards SM, Eden OB, Hall AG, Gibson BE. Analysis of thiopurine methyltransferase variant alleles in childhood acute lymphoblastic leukaemia. *Br J Haematol* 1999;105:696–700. [PubMed: 10354134]
5. Hisamuddin IM, Wehbi MA, Yang VW. Pharmacogenetics and diseases of the colon. *Curr Opin Gastroenterol* 2007;23:60–66. [PubMed: 17133087]
6. Winter JW, Gaffney D, Shapiro D, Spooner RJ, Marinaki AM, Sanderson JD, Mills PR. Assessment of thiopurine methyltransferase enzyme activity is superior to genotype in predicting myelosuppression following azathioprine therapy in patients with inflammatory bowel disease. *Aliment Pharmacol Ther* 2007;25:1069–1077. [PubMed: 17439508]
7. Haglund S, Lindqvist M, Almer S, Peterson C, Taipalensuu J. Pyrosequencing of TPMT alleles in a general Swedish population and in patients with inflammatory bowel disease. *Clin Chem* 2004;50:288–295. [PubMed: 14656901]
8. Lindqvist M, Skoglund K, Karlgren A, Soderkvist P, Peterson C, Kidhall I, Almer S. Explaining TPMT genotype/phenotype discrepancy by haplotyping of TPMT*3A and identification of a novel sequence variant, TPMT*23. *Pharmacogenet Genomics* 2007;17:891–895. [PubMed: 17885628]
9. Colombel JF, Ferrari N, Debuysere H, Marteau P, Gendre JP, Bonaz B, Soule JC, Modigliani R, Touze Y, Catala P, Libersa C, Broly F. Genotypic analysis of thiopurine S-methyltransferase in patients with Crohn's disease and severe myelosuppression during azathioprine therapy. *Gastroenterology* 2000;118:1025–1030. [PubMed: 10833476]
10. Frohman EM, Havrdova E, Levinson B, Slanar O. Azathioprine myelosuppression in multiple sclerosis: characterizing thiopurine methyltransferase polymorphisms. *Mult Scler* 2006;12:108–111. [PubMed: 16459728]
11. Coulthard S, Hogarth L. The thiopurines: an update. *Invest New Drugs* 2005;23:523–532. [PubMed: 16267626]
12. Krynetski EY, Schuetz JD, Galpin AJ, Pui CH, Relling MV, Evans WE. A single point mutation leading to loss of catalytic activity in human thiopurine S-methyltransferase. *Proc Natl Acad Sci U S A* 1995;92:949–953. [PubMed: 7862671]
13. Loennechen T, Yates CR, Fessing MY, Relling MV, Krynetski EY, Evans WE. Isolation of a human thiopurine S-methyltransferase (TPMT) complementary DNA with a single nucleotide transition A719G (TPMT*3C) and its association with loss of TPMT protein and catalytic activity in humans. *Clin Pharmacol Ther* 1998;64:46–51. [PubMed: 9695718]
14. Salavaggione OE, Wang L, Wiepert M, Yee VC, Weinshilboum RM. Thiopurine S-methyltransferase pharmacogenetics: variant allele functional and comparative genomics. *Pharmacogenet Genomics* 2005;15:801–815. [PubMed: 16220112]
15. Hon YY, Fessing MY, Pui CH, Relling MV, Krynetski EY, Evans WE. Polymorphism of the thiopurine S-methyltransferase gene in African-Americans. *Hum Mol Genet* 1999;8:371–376. [PubMed: 9931346]
16. Schaeffeler E, Eichelbaum M, Reinisch W, Zanger UM, Schwab M. Three novel thiopurine S-methyltransferase allelic variants (TPMT*20, *21, *22) - association with decreased enzyme function. *Hum Mutat* 2006;27:976. [PubMed: 16917910]
17. Lindqvist M, Haglund S, Almer S, Peterson C, Taipalensuu J, Hertervig E, Lyrenas E, Soderkvist P. Identification of two novel sequence variants affecting thiopurine methyltransferase enzyme activity. *Pharmacogenetics* 2004;14:261–265. [PubMed: 15083071]
18. Schaeffeler E, Fischer C, Brockmeier D, Wernet D, Moerike K, Eichelbaum M, Zanger UM, Schwab M. Comprehensive analysis of thiopurine S-methyltransferase phenotype-genotype correlation in a large population of German-Caucasians and identification of novel TPMT variants. *Pharmacogenetics* 2004;14:407–417. [PubMed: 15226673]
19. Wang L, Weinshilboum R. Thiopurine S-methyltransferase pharmacogenetics: insights, challenges and future directions. *Oncogene* 2006;25:1629–1638. [PubMed: 16550163]
20. Sasaki T, Goto E, Konno Y, Hiratsuka M, Mizugaki M. Three novel single nucleotide polymorphisms of the human thiopurine S-methyltransferase gene in Japanese individuals. *Drug Metab Pharmacokinet* 2006;21:332–336. [PubMed: 16946561]
21. Krynetski EY, Evans WE. Genetic polymorphism of thiopurine S-methyltransferase: molecular mechanisms and clinical importance. *Pharmacology* 2000;61:136–146. [PubMed: 10971199]

22. Krynetski EY, Evans WE. Pharmacogenetics as a molecular basis for individualized drug therapy: the thiopurine S-methyltransferase paradigm. *Pharm Res* 1999;16:342–349. [PubMed: 10213363]
23. Coulthard SA, Matheson EC, Hall AG, Hogarth LA. The clinical impact of thiopurine methyltransferase polymorphisms on thiopurine treatment. *Nucleosides Nucleotides Nucleic Acids* 2004;23:1385–1391. [PubMed: 15571264]
24. Wusk B, Kullak-Ublick GA, Rammert C, von Eckardstein A, Fried M, Rentsch KM. Thiopurine S-methyltransferase polymorphisms: efficient screening method for patients considering taking thiopurine drugs. *Eur J Clin Pharmacol* 2004;60:5–10. [PubMed: 14985890]
25. Tai HL, Krynetski EY, Yates CR, Loennechen T, Fessing MY, Krynetskaia NF, Evans WE. Thiopurine S-methyltransferase deficiency: two nucleotide transitions define the most prevalent mutant allele associated with loss of catalytic activity in Caucasians. *Am J Hum Genet* 1996;58:694–702. [PubMed: 8644731]
26. Tai HL, Krynetski EY, Schuetz EG, Yanishevski Y, Evans WE. Enhanced proteolysis of thiopurine S-methyltransferase (TPMT) encoded by mutant alleles in humans (TPMT*3A, TPMT*2): mechanisms for the genetic polymorphism of TPMT activity. *Proc Natl Acad Sci U S A* 1997;94:6444–6449. [PubMed: 9177237]
27. Wang L, Nguyen TV, McLaughlin RW, Sikkink LA, Ramirez-Alvarado M, Weinshilboum RM. Human thiopurine S-methyltransferase pharmacogenetics: variant allozyme misfolding and aggresome formation. *Proc Natl Acad Sci U S A* 2005;102:9394–9399. [PubMed: 15967990]
28. Wang L, Sullivan W, Toft D, Weinshilboum R. Thiopurine S-methyltransferase pharmacogenetics: chaperone protein association and allozyme degradation. *Pharmacogenetics* 2003;13:555–564. [PubMed: 12972954]
29. Tai HL, Fessing MY, Bonten EJ, Yanishevsky Y, d'Azzo A, Krynetski EY, Evans WE. Enhanced proteasomal degradation of mutant human thiopurine S-methyltransferase (TPMT) in mammalian cells: mechanism for TPMT protein deficiency inherited by TPMT*2, TPMT*3A, TPMT*3B or TPMT*3C. *Pharmacogenetics* 1999;9:641–650. [PubMed: 10591545]
30. Wu H, Horton JR, Battaile K, Allali-Hassani A, Martin F, Zeng H, Loppnau P, Vedadi M, Bochkarev A, Plotnikov AN, Cheng X. Structural basis of allele variation of human thiopurine-S-methyltransferase. *Proteins* 2007;67:198–208. [PubMed: 17243178]
31. Shield AJ, Thomae BA, Eckloff BW, Wieben ED, Weinshilboum RM. Human catechol O-methyltransferase genetic variation: gene resequencing and functional characterization of variant allozymes. *Mol Psychiatry* 2004;9:151–160. [PubMed: 14966473]
32. Rutherford K, Alphandéry E, McMillan A, Daggett V, Parson WW. Stability of Secondary and Tertiary Protein Structure in 108V and 108M Catechol O-methyltransferase. *BBA*. 2008submitted, minor revision requested, revision returned
33. Cotton NJ, Stoddard B, Parson WW. Oxidative inhibition of human soluble catechol-O-methyltransferase. *J Biol Chem* 2004;279:23710–23718. [PubMed: 15031283]
34. Lotta T, Vidgren J, Tilgmann C, Ulmanen I, Melen K, Julkunen I, Taskinen J. Kinetics of human soluble and membrane-bound catechol O-methyltransferase: a revised mechanism and description of the thermolabile variant of the enzyme. *Biochemistry* 1995;34:4202–4210. [PubMed: 7703232]
35. Rutherford K, Bennion BJ, Parson WW, Daggett V. The 108M polymorph of human catechol O-methyltransferase is prone to deformation at physiological temperatures. *Biochemistry* 2006;45:2178–2188. [PubMed: 16475806]
36. Lee C, Schwartz MP, Prakash S, Iwakura M, Matouschek A. ATP-dependent proteases degrade their substrates by processively unraveling them from the degradation signal. *Mol Cell* 2001;7:627–637. [PubMed: 11463387]
37. Matouschek A. Protein unfolding--an important process in vivo? *Curr Opin Struct Biol* 2003;13:98–109. [PubMed: 12581666]
38. Levitt, M. ENCAD, Energy Calculations and Dynamics. Palo Alto, CA, and Yeda, Rehovot, Israel: Molecular Applications Group; 1990.
39. Beck, DAC.; Alonso, DOV.; Daggett, V. *ilmm, in lucem* Molecular Mechanics. Seattle, WA: University of Washington; 2000–2008.
40. Beck DAC, Daggett V. Methods for molecular dynamics simulations of protein folding/unfolding in solution. *Methods* 2004;34:112–120. [PubMed: 15283920]

41. Levitt M, Hirshberg M, Sharon R, Daggett V. Potential energy function and parameters for simulations of the molecular dynamics of proteins and nucleic acids in solution. *Computer Physics Communications* 1995;91:215–231.
42. Levitt M, Hirshberg M, Sharon R, Laidig KE, Daggett V. Calibration and testing of a water model for simulation of the molecular dynamics of proteins and nucleic acids in solution. *Journal of Physical Chemistry B* 1997;101:5051–5061.
43. Kell GS. Precise representation of volume properties of water at one atmosphere. *Journal of Chemical Engineering Data* 1967;12:66–69.
44. Beck DA, Armen RS, Daggett V. Cutoff size need not strongly influence molecular dynamics results for solvated polypeptides. *Biochemistry* 2005;44:609–616. [PubMed: 15641786]
45. Hubbard, SJ.; Thorton, JM. NACCESS. London: Department of Biochemistry and Molecular Biology, University College; 1993.
46. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* 2004;25:1605–1612. [PubMed: 15264254]

Figure 1. Thiopurine *S*-methyltransferase polymorphisms

(a) Ribbon diagram of TPMT (2BZG.PDB, 30) colored from blue (N-terminus) to red (C-terminus). Polymorphic residues 80, 154 and 240 are shown in space-filling representation and colored by atom. The sulfur atom of SAH points towards residues F40 and P196, which define a pore on the protein's surface. Co-product *S*-adenosylhomocysteine (SAH) and solvent molecule 2-[3-(2-hydroxy-1,1-dihydroxymethyl-ethylamino)-propylamino]-2-hydroxymethyl-propane-1,3-diol (B3P) are shown in licorice representation and colored by atom. (b) Space-filling structure of TPMT. The thiopurine-binding site is composed of residues from α1, α6, α9 and the loop between β7 and α8. F40 and P196 border a water channel on the protein's surface that allows thiopurine substrates to bind and interact with SAM. Figure is

rotated 90° from structure shown in 1a. SAH, F40 and P196 are colored by atom. The rest of the protein is colored to match Figure 1a.

Figure 2. Mobility of the TPMT variants at 37°C

The figure shows C α RMSF (Å) plots of the wild-type (*TPMT* *1) protein compared with *TPMT* (a) *2 (A80P), (b) *3B (A154T), (c) *3C (Y240C), (d) *3A (A154T, Y240C). C α RMSF values for the wild-type and polymorphic proteins are colored in blue and red, respectively. Crystallographic B-factors from the wild-type structure (2BZG, PDB, ³⁰) bound with SAH and B3P are colored in black. α -helices and β -sheets are denoted by bars and arrows, respectively, and colored to match the structures in Figure 1. * polymorphic residue. ↓ large fluctuations.

Figure 3. Effects of TPMT polymorphisms on local tertiary structure

(a, b) A80 forms side-chain contacts with residues in $\alpha 3$ (K77, D81, R82), $\beta 2$ (V86), and strand $\beta 5$ (N127, I128). The A80P polymorphism breaks the intra-helical 80 N – M76 O hydrogen bond, disrupting the last turn of $\alpha 3$. This results in the loss of the P80 interaction with K77 as well as with several residues β -strand. (c, d) The hydrophobic C β atoms of both A154 and T154 face the interior of the SAM-binding site. However, T154 O γ forms an additional hydrogen bond with D151, disrupting the Y166 ($\alpha 7$) – D151 ($\alpha 6$)– Y180 ($\beta 7$) network and opening up the SAM-binding site. (e, f) Y240 interacts with residues in $\beta 7$ (L182, V184), $\beta 8$ (I214, C216, L217), $\beta 9$ (K238, L242) and $\alpha 8$ (H201, I204, E205). The smaller C240 maintains most of the β -strand contacts, but loses most interactions with residues in $\alpha 8$. Snapshots from

the 20th ns of three independent 37°C simulations are superposed in c, d, e, f. Secondary structural elements are colored as in Figure 1. Polymorphic residues and nearby side-chains are shown in space-filling and licorice representation, respectively, and colored by atom.

Figure 4. The *TPMT2 allele affects local salt-bridge network**

(a) The loss of $\alpha 3$ disrupts the nearby D54-R82 salt-bridge and stabilizes/creates a K77-E102 salt bridge. The change in packing alters the orientations of $\alpha 2$, $\alpha 3$, $\alpha 4$ and $\beta 2$, which have SAM-binding residues at their distal ends. Residues in the salt-bridge network and SAM-binding site are shown in space-filling and licorice representations, respectively, and colored by atom. Residue 80 is colored in magenta. (b) The A80P polymorphism alters the SAM-binding site. The reorientation of $\alpha 2$ pulls $\alpha 1$ away from the SAM-binding site while the additional K77-E102 salt bridge tightens the interaction between $\alpha 3$ and $\alpha 4$, moving $\alpha 4$ away from the active site. These movements result in the formation of a large cleft in the protein. $\alpha 1$, $\alpha 3$, and $\alpha 4$ are colored in blue, cyan and green respectively. Residue 80 is shown in space-filling representation and colored magenta.

Figure 5. The *TPMT* *3A allele disrupts the overall structure of the proteins

Initial and 20 ns structures of (a) *TPMT* *1 (WT), (b) *TPMT* *3C, and (c) *TPMT* *3A are shown in ribbon diagram. *TPMT* *3A flattens and expands during the simulations. Despite being positioned quite far from each other, polymorphic residues 154 and 240 are connected through a hydrogen-bonding network. T154-D151-Y180 form side-chain interactions, while β -strands 7 and 9 form backbone hydrogen bonds between Y180-L242 and L182-C240. These interactions are maintained throughout the simulations. (b) The A154T polymorphism opens up the SAM-binding site, while the Y240C polymorphism loses contacts with residues in helix α 8. This loosening up of the C-terminus facilitates the rearrangement of the β 7- α 8 loop and the opening up of the thiopurine-binding site. Side-chains are shown in space-filling mode and colored by atom.

Figure 6. Potential sites of TPMT ubiquitination

TPMT has 21 lysine residues distributed throughout the protein surface. In our MD simulations, 5 lysine residues (K20, K37, K60, K145, K245) located in unstructured regions of the protein become more solvent exposed in the *TPMT**2A, *3B, *3C, and *3A simulations. Any of these lysines may serve as substrates for ubiquitination. Ribbon diagram of *TPMT**3A from the 20th ns of a 37°C simulation is colored 30 from blue (N-terminus) to red (C-terminus). Lysine residues and polymorphic residues A154T and Y240C are shown in stick and space filling representations, respectively, and colored by atom.

Figure 7. TPMT polymorphisms

The TPMT sequence contains 23 polymorphisms distributed throughout the protein. Four of the polymorphic residues (Y180F, R215H, C216R, K238E) are clustered around Y240C in the C-terminal β -strands 7, 8, and 9. Three polymorphic residues are located in the SAM (L69V) and (Q42E, H227Q) thiopurine binding sites. SAM and the polymorphic residues are colored by atom and shown in space-filling and licorice representations, respectively.

Table 1General Properties of the *TPMT* *2 (A80P) Simulations

General Properties ^a	Initial (WT, 2BZG)	<i>TPMT</i> *1 (WT)	<i>TPMT</i> *2 (A80P)
SASA: Total (Å ²)	10913	12466 ± 313	12596 ± 251
SASA: SAM Site (Å ²) ^b	425	561 ± 127	783 ± 97 ^f
SASA: Thiopurine Site (Å ²) ^c	670	521 ± 79	584 ± 54
Cα RMSD of Helix α3 (Å)	NA	0.4 ± 0.1	1.7 ± 0.2 ^e
Polymorphic Site			
SASA: Polymorphic Site (Å ²) ^d	310	396 ± 43	387 ± 21
d(A80 CB/ P80 CD – K77 CD) (Å)	4.5	6.2 ± 1.0	9.5 ± 0.9 ^g
d(A80 CA/ P80 CB – G83 CA) (Å)	5.2	4.9 ± 0.3	5.7 ± 1.6
d(A80 CA/P80 CA – H84 CD2) (Å)	6.1	6.5 ± 0.4	8.7 ± 2.2
d(A80 CB /P80 CD – L105 CD2) (Å)	7.7	7.5 ± 1.0	6.9 ± 1.7
d(A80 CB/P80 CG – S125 CB) (Å)	7.4	9.0 ± 2.7	7.3 ± 2.1
d(A80 CB/P80 CG – N127 CB) (Å)	4.3	6.2 ± 2.7	4.8 ± 0.7
d(A80 CB/P80 CD – I128 CG1) (Å)	4.0	5.1 ± 1.6	5.3 ± 0.8
Salt Bridges			
d(D54 OD2 – R82 NH2) (Å)	3.4	3.3 ± 0.7	10.1 ± 3.3 ^g
d(K77 NZ – E102 OE1) (Å)	13.1	9.7 ± 3.8	3.0 ± 0.8 ^g
d(E75 OE2 – R152 NH2) (Å)	6.2	4.7 ± 1.3	6.3 ± 1.0

^a All properties were calculated using structures from the last 5 ns (5,000 structures) of each simulation.

^b The SAM-binding site was defined using the following residues: L26, W29, W33, F40, L69, G71, E90, I91, S134, I135, R152, G153, A157.

^c The thiopurine-binding site was defined using the following residues: W29, K32, A39, F40, Q42, R152, V156, A157, C183, L185, P195, P196, A222, E224, R226, H227, W230 and W237.

^d The A80P polymorphic site was defined using the following residues: M76, K77, F79, D81, R82, G83, H84, V86, S125, N127, I128 and residue 80. NA: not applicable.

^e p < 0.005

^f p < 0.10

^g p < 0.05.

Table 2

Structural Effects of the A154T and Y240C TPMT Polymorphisms

General Properties ^a	TPMT*1 (WT)	TPMT*3B (A154T)	TPMT*3C (Y240C)	TPMT*3A (A154T/Y240C)
SASA: Total (Å ²)	12466 ± 313	12873 ± 571	12522 ± 252	12586 ± 1554
SASA: SAM Site (Å ²) ^b	561 ± 127	628 ± 129	508 ± 105	567 ± 190
SASA: Thiopurine Site (Å ²) ^c	521 ± 79	653 ± 117	628 ± 62	790 ± 115 ^g
Cα RMSD of α1 (Å)	3.5 ± 1.1	3.7 ± 0.5	3.0 ± 0.7	3.4 ± 1.1
Cα RMSD of α5 (Å)	1.8 ± 0.8	0.7 ± 0.1 ^f	1.6 ± 0.8	0.6 ± 0.3 ⁱ
Cα RMSD of α7 (Å)	0.7 ± 0.4	1.2 ± 0.2	0.5 ± 0.2	0.8 ± 0.4
Cα RMSD of α8 (Å)	0.4 ± 0.1	0.4 ± 0.1	0.6 ± 0.2	0.5 ± 0.2
Cα RMSD of α9 (Å)	1.2 ± 0.3	1.5 ± 0.2	1.7 ± 0.2	1.8 ± 0.4
A154T Polymorphic Site				
SASA: Polymorphic Site (Å ²) ^d	143 ± 73	274 ± 127	140 ± 37	169 ± 104
SASA: Residue 154 (Å ²)	5.5 ± 6.2	17.1 ± 14.4	9.1 ± 7.6	11.1 ± 10.5
d(A154 CB/T154 CG2 – P68 CB) (Å)	4.6 ± 0.5	5.2 ± 1.6	5.0 ± 1.1	4.6 ± 0.5
d(A154 CB/T154 CG2 – L69 CB) (Å)	6.4 ± 1.5	7.2 ± 0.6	6.2 ± 1.3	6.0 ± 1.5
d(A154 CB/T154 CG2 – I135 CD1) (Å)	10.7 ± 5.8	11.0 ± 5.0	8.9 ± 3.3	8.3 ± 3.5
d(A154 CB/T154 CG2 – Y166 CE1) (Å)	4.3 ± 0.4 ^l	8.1 ± 2.5 ^f	4.9 ± 1.0	5.5 ± 0.6
d(D151 OD2 – Y166 OH) (Å)	3.0 ± 0.9 ^k	11.4 ± 4.5 ^g	3.5 ± 1.8 ^k	5.6 ± 1.3 ^k
d(D151 OD2 – Y180 OH) (Å)	2.5 ± 0.1	2.5 ± 0.1	2.5 ± 0.1	2.5 ± 0.1
d(Y166 OH – Y180 OH) (Å)	3.5 ± 0.4	11.2 ± 5.2 ^f	4.1 ± 1.4 ^k	4.8 ± 1.2 ^k
Y240C Polymorphic Site				
SASA: Polymorphic Site (Å ²) ^e	448 ± 15	531 ± 12 ^h	485 ± 65	555 ± 88 ^f
d(Y240 OH/C240 SG – H201 ND1) (Å)	5.5 ± 1.3	10.4 ± 2.1 ^g	7.6 ± 1.7	9.4 ± 3.1 ^f
d(Y240 CB/C240 SG – I204 CG2) (Å)	7.7 ± 1.3	7.1 ± 1.3	6.7 ± 0.8	7.7 ± 1.2
d(Y240 OH/C240 SG – E205 OE2) (Å)	7.3 ± 4.5	11.4 ± 1.3	11.7 ± 1.2	13.1 ± 1.7 ^j

^a All properties were calculated using structures from the last 5 ns (5,000 structures) of each simulation.^b The SAM-binding site was defined using the following residues: L26, W29, W33, F40, L69, G71, E90, I91, S134, I135, R152, G153, A157.^c The thiopurine-binding site was defined using the following residues: W29, K32, A39, F40, Q42, R152, V156, A157, C183, L185, P195, P196, A222, E224, R226, H227, W230 and W237.^d Residues P68, L69, I135, D151, G153, L155, V156, A156, I157, Y166 and residue 154 defined the A154T polymorphic site.^e Residues L182, V184, H201, I204, E205, I214, C216, L217, K219, K238, L239, L242 and residue 240 defined the Y240C polymorphic site.^f $p < 0.1$ to TPMT*1.^g $p < 0.05$ to TPMT*1.^h $p < 0.01$ to TPMT*1.ⁱ $p < 0.1$ to TPMT*3C.^j $p < 0.05$ to TPMT*3C.^k $p < 0.1$ to TPMT*3B.^l $p < 0.05$ to TPMT*3B.