Published in final edited form as:

Trends Biochem Sci. 2010 October ; 35(10): 531–538. doi:10.1016/j.tibs.2010.03.007.

Polymorphisms and Disease: Hotspots of Inactivation in Methyltransferases

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

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

[‡] Department of Bioengineering, Box 355103, University of Washington, Seattle WA 98195-5013

Abstract

Methyltransferases catalyze methylation processes essential for protein/DNA repair, transcriptional regulation, and drug metabolism *in vivo*. More than 500 human methyltransferase polymorphisms have been identified, many of which are linked to disease. We mapped all available coding polymorphisms of seven methyltransferases onto their structures to address their structural significance, and identified a polymorphic ‘hotspot’ ~20 Å from the active site in four of the proteins. Molecular dynamics simulations of these proteins reveal a common mechanism of destabilization: the mutations alter important side-chain contacts within the polymorphic site that are propagated through the protein, thereby distorting the active site. We propose that this ‘hotspot’ might have arisen to modulate enzymatic activity, with decreased activity actually conferring an advantage in three of the four methyltransferases.

An Abundance of Methyltransferase Polymorphisms

Single nucleotide polymorphisms (SNPs) are the most common source of human genetic variation, occurring at a frequency of 1 in every 1000 nucleotide bases of the 3 billion-base human genome¹. SNPs can be intronic or exonic and can then affect mRNA transcription, structure, splicing and translation, protein expression, structure, enzymatic activity and stability. The completion of the Human Genome Project in 2003 gave way to the development of several publicly accessible databases listing millions of SNPs. In 2007, the International HapMap Project published a study identifying more than 3.1 million SNPs in 270 individuals². The idea that this astronomical source of genetic variation is linked with a comparably large number of genotypic defects has piqued interest into clinical associations between specific haplotypes and disease. However, functional studies describing the biochemical effects of these polymorphisms lag quite far behind the cataloging of these genetic variations.

Methyltransferases catalyze the methylation of proteins, RNA, and DNA, playing an essential role in protein repair, epigenetic silencing, cell mobility, cell adhesion, neurotransmission and drug and hormone metabolism³. Interest in the genetic variation of methyltransferase proteins has increased significantly over the past 10 years. Indeed, over 500 polymorphisms have been identified in human methyltransferase genes, and several are associated with increased risk for

*Corresponding Author: Phone: (206) 685-7420, daggett@u.washington.edu.

§Present address: Department of Biochemistry and Biophysics, Howard Hughes Medical Institute, University of Pennsylvania School of Medicine, Philadelphia PA 19104

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.

disease (Table S1). Large-scale functional genomics studies have characterized how coding polymorphisms in catechol *O*-methyltransferase (COMT)⁴, histamine *N*-methyltransferase (HNMT)⁵, protein *L*-isoaspartate *O*-methyltransferase (PIMT)⁶, thiopurine *S*-methyltransferase (TPMT)⁷, glycine *N*-methyltransferase (GNMT)^{8,9}, guanidinoacetate *N*-methyltransferase (GAMT)¹⁰, phenylethanolamine *N*-methyltransferase (PNMT)¹¹, phosphatidylethanolamine *N*-methyltransferase (PEMT)¹², and arsenic methyltransferase (AS3MT)¹³ affect protein activity and stability. However, the molecular bases for these mutation-induced effects remain to be determined. Here we have identified a polymorphic hotspot in COMT, HNMT, PIMT, and TPMT located ~20Å from the active site at the intersection of $\alpha 2$, $\alpha 3$, and $\beta 3$. Substitutions at this hotspot effect structural rearrangements of nearby helices resulting in a distortion of the *S*-adenosylmethionine (SAM) binding site. These structural changes might account for the altered enzymatic activity associated with each polymorphism.

Methyltransferase Pharmacogenetics

COMT was one of the first proteins shown to display large inter-individual differences in enzymatic activity. COMT activity is distributed in a trimodal pattern with low (LL), intermediate (LH), and high (HH) activity, where individuals homozygous for LL have two to four-fold lower enzyme activity than those homozygous for the HH allele¹⁴. Restriction fragment length polymorphism mapping revealed that these differences in activity were attributable to a valine to methionine substitution at residue 10815. The V108M variant results in decreased protein levels *in vivo*¹⁶, and epidemiological studies have associated the V108M substitution with increased risk for breast cancer¹⁷ and neuropsychiatric diseases¹⁸ (Table S1).

TPMT, which is responsible for the metabolism of cytotoxic thiopurine drugs used to treat a variety of conditions including childhood leukemia, Crohn's Disease, and inflammatory bowel disease¹⁹, has recently emerged an important pharmacogenetic target. Although the thiopurine drugs are effective, they have a relatively narrow therapeutic index and prolonged exposure can be fatal. More than 20 *TPMT* polymorphisms resulting in intermediate to no *TPMT* activity have been identified in 11% of the population. Owing to this extreme inter-individual variation in *TPMT* activity, genotyping and activity measurements are now performed routinely prior to thiopurine treatment¹⁹.

Crystallographic and NMR studies of the polymorphic variants are critical to understanding how the underlying polymorphisms affect protein structure and dynamics. Unfortunately, however, many of the variant forms are difficult to characterize structurally due to significantly increased rates of protein degradation and aggregation. Another frequent problem is that the average, static structures of paired wild-type and polymorphic proteins are typically very similar and fail to shed light on the effect of mutation. For example, the available crystal structures of variant COMT and HNMT proteins have been achieved only in the presence of stabilizing substrate molecules, providing little insight into the effects of each polymorphism.²⁰⁻²¹ By contrast, molecular dynamics (MD) studies are not plagued by aggregation or stability issues, and they can provide high-resolution information regarding protein structures, dynamics, and excursions from the average conformation in solution. In these instances, MD simulations of the wild-type and polymorphic proteins can both complement and serve as surrogates for experimental studies by providing important clues regarding how the mutations affect protein structure and dynamics at an atomic level.

We have mapped 18 non-synonymous coding polymorphisms whose products decrease protein activity onto the available crystal structures of the respective human methyltransferase (Figure 1A). The substitutions are not localized to the active site, but instead are distributed throughout each protein making it difficult to predict how they might affect methyltransferase structure.

MD studies of COMT, HNMT, PIMT, and TPMT detail the overall structural and dynamic effects of each polymorphism. Cataloguing the molecular bases underlying the destabilizing effects of polymorphisms in these structurally and biochemically well-characterized methyltransferases can provide a solid basis for predicting the structural impact of novel polymorphisms in less studied protein systems.

A polymorphic hotspot in four methyltransferase proteins

Our studies uncovered a ‘hotspot’ for polymorphic variation at the intersection of $\alpha 2$, $\alpha 3$, and $\beta 3$ in the structures of COMT, HNMT, PIMT, and TPMT. The V108M (COMT) and T105I (HNMT) substitutions are positioned in the surface loop between $\alpha 3$ and $\beta 3$ where the side-chain of each variant residue is buried within a hydrophobic pocket comprising residues from $\alpha 2$, $\alpha 3$, and $\beta 3$ (Figure 1B). The V119I substitution of PIMT is located in the second turn of helix $\alpha 3$. Interestingly, residue 119 forms van der Waals contacts with L130, which is located at a position in the $\alpha 3$ - $\beta 3$ loop identical to that of residues 108 of COMT and 105 of HNMT (Figure 1A). The A80P substitution of TPMT is located in the last turn of helix $\alpha 2$, where it interacts with N127 and I128 in $\beta 3$ (Figure 1B). Each of these seemingly innocuous substitutions are located ~16–20 Å from the protein’s active site. However, α -helices 1–4 and β -strands 1–3 all contain SAM-binding residues on their distal ends – most notably a conserved acidic residue (for example, E89, E90, and D109) that forms hydrogen-bonds with the ribose hydroxyl groups of SAM (Figure 1B). Structural changes at the polymorphic site could thus be relayed to the active site. Indeed, each variant displays decreased protein activity *in vitro*^{16, 20, 22, 23} (Table S1). In addition, the COMT (V108M) and TPMT (A80P) substitutions destabilize the structure of the protein and decrease protein levels *in vivo*^{16, 23–25}. To examine how mutations in this ‘polymorphic hotspot’ affect active-site structure and dynamics, multiple MD simulations of the wild-type and variant COMT, HNMT, PIMT, and TPMT apoprotein structures at 37°C were performed.

The COMT (V108M) and HNMT (T105I) variants have very similar effects on active-site structure during the MD simulations^{26, 27}, possibly due to their nearly identical positions in the COMT and HNMT protein structures. Substitution of the larger Met for Val in COMT and Ile for Thr in HNMT tightens packing interactions between M108/I105 and residues in $\alpha 2$, $\alpha 3$ and $\beta 3$, effectively burying the variant residue (Figure 2). This repacking distorts $\alpha 4$, while increasing the flexibility of active-site residues in $\alpha 3$, $\beta 3$, and $\alpha 4$ ^{26, 27} of both HNMT and COMT, and increasing the solvent exposure of the COMT SAM-binding site by ~200 Å²²⁷ (Figure 2). These changes are translated throughout the protein, affecting the orientation of $\alpha 6$ (COMT) and αA (HNMT) and disrupting the substrate-binding site. The wild-type and V108M COMT variants show similar kinetic properties *in vitro*. However, V108M COMT is 2–4 fold less active *in vivo* due to more rapid proteolytic degradation and decreased levels of immunological protein^{16, 28}. SAM has been shown to increase both the chemical and thermal stability of V108M COMT²⁵. Both COMT variants remain folded throughout the simulations. However, the distortion of $\beta 3$ and $\alpha 4$, which border the SAM site, might represent a nucleation site for the destabilization of V108M COMT.

PIMT displayed the subtlest structural effects upon substitution in the hotspot region. Both V119 and I119 maintain contacts with the same residues in $\alpha 3$ and $\beta 3$ throughout the simulations. However, the larger Ile disrupts sidechain packing between residues in $\alpha 3$ and $\beta 3$, reorienting $\beta 3$, $\alpha 3$ and $\alpha 4$ (Figure 2). The displacement of $\beta 3$ and $\alpha 4$ increases the mobility of several residues (H110, D141, R143) in the SAM-binding site increasing its solvent exposure by 200 Å²²⁹ and perhaps facilitating SAM binding (Figure 2). Because the SAM- and isoaspartate-binding sites share several common residues (I52, Q221), this disruption of the active site is translated to the isoaspartate interface on the opposite face of the protein. In contrast, the I119V variant remains very compact during the simulations with the SAM site

buried by α C and inaccessible, and the isoaspartate site intact (Figure 2). These subtle differences provide insight into the observation that although both PIMT variants are folded, active proteins, the I119V protein shows both decreased overall activity and an increased specificity for endogenous substrates over wild-type PIMT^{6, 22}.

The TPMT A80P variant is proteolytically degraded with a half-life of ~0.2 hrs, resulting in very low levels of immunological protein *in vivo* and negligible TPMT activity^{23, 30}. Substituting a helix-breaking Pro residue for Ala at position 80 completely disrupts the final turn of helix α 2, resulting in a 1.3 Å reorientation of α 2 which pulls residue 80 away from residues in α 2 and β 2 while breaking several salt-bridge networks which connect helices α 1 through α 3³¹ (Figure 2). Because α 1, α 2, β 2, and β 3 all contain SAM-binding residues, the A80P induced disruption of their local environment has a large impact on active-site structure. Indeed, a large cleft forms in the protein, deforming the active-site and increasing the solvent exposure of the SAM-binding site by ~200 Å²³¹ (Figure 2). These destabilizing structural changes might nucleate the rapid protein degradation caused by the A80P substitution accounting for the complete loss of TPMT activity^{23, 30}.

The simulations reveal that the COMT, HNMT, and PIMT substitutions appear to alter protein structure via a common molecular mechanism. The introduction of a larger side chain (M108, I105, I119) in this polymorphic hotspot buries the variant residue within the polymorphic site and causes the rearrangement of α 2, α 3, β 3, and α 4, all of which contain SAM-binding residues on their distal ends. The increased flexibility of the active-site residues loosens up the SAM-binding site, increasing its solvent exposure. Interestingly, intermediate COMT, HNMT, and PIMT activity appears to be advantageous. Indeed, whereas the decreased ability of V108M COMT to metabolize estrogen and norepinephrine is associated with increased risk for breast cancer³² and neuropsychiatric disease³³, it has also been linked with improved prefrontal cognition, especially in working memory³⁴. Similarly, elevated levels of histamine in individuals homozygous for the allele encoding HNMT 105I provides protection against infectious agents and affects anxiety levels, cognition and sedation³⁵. Heterozygosity resulting in the expression of both PIMT 119V and 119I variants favors the efficient repair of isoaspartate residues by allowing both the fast (119I) and specific (119V) methylation of damaged proteins²². It is an intriguing possibility that this ‘polymorphic hotspot’ arose as a means to modulate enzymatic activity, with the effects of ‘innocuous’ (COMT, HNMT, PIMT) and ‘detrimental’ (TPMT) substitutions ranging from moderate decreases in activity to rapid degradation of the protein.

Several Attractive Polymorphic Methyltransferase Targets Remain

The relative abundance of structural, biochemical, and epidemiological data available for the wild-type and variant COMT, HNMT, PIMT, and TPMT proteins made these systems ideal candidates for study via MD simulations. However, structural and/or functional information is lacking for how polymorphisms affect GNMT, GAMT, PNMT, PEMT, and AS3MT structure and dynamics, making validation of the MD simulations more difficult.

GNMT constitutes ~1–3% of all cytosolic hepatic proteins and is a key enzyme involved in maintaining the cell’s methylation (SAM:S-adenosylhomocysteine(SAH)) capacity³⁶. GNMT deficiency was first identified in individuals with hepatomegaly and mild liver disease who also displayed hypermethioninaemia^{8, 37}. Re-sequencing of *GNMT* identified three polymorphisms (encoding L49P, N140S, and H176N) that decrease GNMT activity^{8, 37, 38}. Several crystal structures of the rat and human GNMT homotetramer are available, including a structure of the H176N variant³⁹. Although the static structures can be used to predict the detrimental effects of the L49P and N140S substitutions on structure and activity, the destabilization of GNMT by H176N is more difficult to discern (Box 1). The extensive

biochemical and structural characterization of the GNMT polymorphisms makes GNMT an intriguing candidate for MD studies. However, missing electron density from large portions of the GNMT crystal structures, coupled with the large size of the active tetrameric protein complex (130 kD) currently makes these studies challenging.

Box 1

GNMT Polymorphisms Alter overall Protein Structure: A Future Target

GNMT is active only as a homotetrameric protein. A series of hydrogen bonds and salt bridges between helices $\alpha 4$ of two monomers define the dimer interface. The homotetramer is formed through antiparallel hydrogen-bonding interactions between the symmetry related β -sheets of two dimers (Figure IA). Interestingly, each flexible N-terminus contributes to the active sites of both its respective monomer and that of the adjacent subunit, and its position within the active site is critical for both the catalytic activity of GNMT and for substrate-product exchange⁵³.

Several deactivating polymorphisms are distributed throughout the GNMT structure. The L49P substitution reduces GNMT activity by 90%. L49P also decreases the protein's K_m for SAM, despite being positioned in the second turn of $\alpha 1$, ~20 Å from the SAM-binding site (Figure IA)⁹. This substitution does not affect GNMT quaternary or tertiary structure, but does decrease GNMT's overall helical content⁹. Helices $\alpha 1$ and αA are linearly connected by a small unstructured region and any loss of $\alpha 1$ helicity due to a helix-breaking proline residue could be propagated through αA to the N-terminus. The importance of the N-terminus suggests that any L49P-induced rearrangement of this region could ablate GNMT activity.

The N140S substitution has no detectable impact on GNMT structure. However, it completely ablates GNMT activity while increasing GNMT's K_m for glycine⁵⁴. Residue 140 is located within the SAM- and glycine-binding sites where it forms hydrophobic interactions with SAM and a N140-C γ -Gly-O1 hydrogen bond with glycine (Figure I). It is probable that the smaller serine residue disrupts these interactions, affecting active-site structure and, therefore, enzymatic activity.

The H176N substitution decreases GNMT's activity and affinity for glycine by 25% and 2-fold, respectively, probably by destabilizing both the tetrameric and monomeric structure of GNMT^{39, 54}. Although a crystal structure of the H176N GNMT variant has been solved³⁹, the mechanism by which this substitution effects these changes remains largely unknown. Residue 176 is positioned ~29 and 22 Å from the dimer and tetramer interfaces, respectively (Figure IA), in the active-site periphery adjacent to R77, a key residue in glycine binding (Figure IB). N176 is twisted relative to H176, allowing the N176-OD1 and H288-NE2 atoms to come within 3 Å of each other. However, this additional hydrogen bond does not affect the orientations of nearby catalytic residues. The extensive biochemical characterization of the H176N substitution, along with this inconclusive structural data, makes GNMT an intriguing candidate for MD studies.

PNMT catalyzes the methylation of norepinephrine to form epinephrine^{40, 41}. Re-sequencing of *PNMT* identified four polymorphisms (encoding N9S, T98A, R112C, and A175T); the T98A substitution significantly decreases the levels of PNMT enzymatic activity and immunoreactive protein (Table S1)^{11, 42}. Although there are no clinical reports linking the expression of the T98A variant of PNMT with disease, two non-coding polymorphisms flanking the 5' end of the *PNMT* gene have been associated with increased risk for hypertension⁴³, sporadic early onset Alzheimer's disease⁴⁴, and multiple sclerosis⁴⁵. Several crystal structures of PNMT are available, along with detailed studies regarding the mode of

substrate binding and catalysis, making the T98A PNMT mutant an interesting candidate for structural and dynamic studies.

GAMT is responsible for the biosynthesis of creatine, using ~ 75 % of the total cellular SAM stores in the process⁴⁶. Several amino acid substitutions (W20S, H51P, A53P, L197P) are associated with GAMT deficiency, a disorder characterized by the accumulation of guanidinoacetate and the depletion of creatine in the brain which leads to early onset epilepsy and mental retardation (Table S1)^{10, 47}. Crystal structures of the human and rat GAMT proteins, along with detailed mutagenic studies describing the catalytic mechanism of GAMT, are available. In addition, one study showed that the L197P substitution markedly reduced GAMT activity⁴⁷. However, elucidating the effects of these mutations on GAMT activity has been extremely difficult as most of these mutations have been identified only in single individuals, usually in conjunction with mutations in either the intronic or 3' untranslated regions (UTRs). Additional biochemical and clinical studies tracking the occurrence of these mutations would make GAMT an attractive candidate for MD simulations.

Coding polymorphisms in *AS3MT*¹³, and *PEMT*⁴² affect the enzymatic activity of their protein products and have been linked with increased risk for the development of arsenic-induced skin lesions and cancer⁴⁸, and nonalcoholic fatty liver disease⁴⁹, respectively. However, to date there are no high-resolution structures available for AS3MT or PEMT.

Concluding remarks

A significant number of non-synonymous coding polymorphisms affect both the enzymatic activity and the overall stability of methyltransferase proteins. We have identified a 'polymorphic hotspot' in four methyltransferase proteins [COMT (V108M), HNMT (T105I), PIMT (I119V), and TPMT (A80P)] ~16–20 Å from the SAM-binding site. All of these polymorphisms decrease protein activity, and the TPMT polymorphism drastically reduces levels of immunological protein. MD studies of the wild-type and variant proteins suggest a common mechanism of destabilization in which seemingly innocuous COMT, HNMT, and PIMT amino acid substitutions alter important sidechain contacts within the polymorphic site, inducing the rearrangement of neighboring $\alpha 3$, $\alpha 4$ and $\beta 3$. These changes are propagated through the protein structure, distorting the positions of several key SAM-binding residues and altering the overall solvent exposure of the SAM- and substrate-binding sites. The structural effects of the A80P TPMT substitution are more severe as it breaks a helix, opening up a large cleft in the protein that could nucleate protein degradation. The MD studies of COMT, HNMT, PIMT, and TPMT provide structural and dynamic information critical to mapping the effects of the underlying mutations. This information potentially can be used to develop molecules to stabilize the structure as well as protein activators for the treatment of disease. Cataloguing the molecular bases underlying the destabilizing effects of polymorphisms in the experimentally well characterized methyltransferases should provide fodder for predicting the structural impact of novel polymorphisms in less studied protein systems.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

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). Molecular graphics images were produced using the UCSF Chimera package⁵⁰.

Glossary Box

$\text{C}\alpha$ -root-mean-squared deviation ($\text{C}\alpha$ -rmsd)	root mean square deviation of $\text{C}\alpha$ atoms in a protein structure at a time point during a MD simulation from those of the starting structure. $\text{C}\alpha$ -rmsd is a measure of how the structure is changing over time
Methyltransferase	protein that catalyzes the transfer of a methyl group from co-substrate S-adenosylmethionine (SAM) to an O, N or S atom of a substrate (protein, DNA, RNA, lipid, small molecule). Class I SAM-dependent methyltransferases have a conserved mixed α/β Rossmann-like fold consisting of a core seven-stranded β -sheet (ordered $3\uparrow 2\uparrow 1\uparrow 4\uparrow 5\uparrow 7\downarrow 6\uparrow$) sandwiched between two layers of α -helices (Figure 1). SAM interacts with conserved motifs at the C-terminal ends of β -strands 3, 2, 1 and 4. Substrate binding and specificity are determined by additional structural elements at the protein's amino- and carboxy-termini. Approximately 150 proteins have been characterized as methyltransferases and over 650 structures have been deposited in the protein data bank
Molecular Dynamics (MD)	computational methodology that describes molecular motions by calculating the changes of atomic coordinates as a function of time. MD simulations of proteins require a high-resolution crystal or NMR structure of the protein to be studied to serve as a static starting structure. Any missing atoms are then added (crystal structures, for example, lack hydrogen atoms) and the protein is solvated with water molecules in a periodic box where the solvent density is set to the experimentally determined value for the desired temperature of the run. The atom velocities are then slowly increased to values corresponding to those of the desired temperature. The new atom positions and velocities are calculated by solving the equations of motion using values of the previous atom positions, accelerations and velocities. Each calculation is performed using a very small time step (fs) in order to accurately capture the structural and dynamical changes of the system
Pharmacogenetics	relationship between genetic inheritance and variations in drug dosing or drug response phenotypes. The effects can be life threatening if the genotype in question affects drug metabolism leading to the accumulation of toxic intermediates. For example, polymorphisms in thiopurine S-methyltransferase (<i>TPMT</i>) can lead to variations in the metabolism of the thiopurine drugs used to treat childhood leukemia. Patients are now genotyped prior to dosing to prevent adverse effects
Polymorphic Site	amino acids that surround and interact with the SNP-associated residue in a protein
Polymorphism	genetic variation that occurs with a frequency of >1% within the general population
S-AdenosylMethionine (SAM)	the second most abundant co-factor in the body after ATP. SAM comprises an adenosine and methionine moiety and is a key co-substrate for methyl transfer reactions, where the methyl group on the SD atom of SAM's methionine moiety is transferred to a

substrate yielding a methylated product and S-adenosylhomocysteine (SAH)

References

1. Pennisi E. A closer look at SNPs suggests difficulties. *Science* 1998;281:1787–1789. [PubMed: 9776677]
2. Consortium TIH. A second generation human haplotype map of over 3.1 million SNPs. *Nature* 2007;449:851–861. [PubMed: 17943122]
3. Loenen WA. S-adenosylmethionine: jack of all trades and master of everything? *Biochem Soc Trans* 2006;34:330–333. [PubMed: 16545107]
4. Shield AJ, et al. Human catechol O-methyltransferase genetic variation: gene resequencing and functional characterization of variant allozymes. *Mol Psychiatry* 2004;9:151–160. [PubMed: 14966473]
5. Preuss CV, et al. Human histamine N-methyltransferase pharmacogenetics: common genetic polymorphisms that alter activity. *Mol Pharmacol* 1998;53:708–717. [PubMed: 9547362]
6. David CL, et al. Human erythrocyte protein L-isoaspartyl methyltransferase: heritability of basal activity and genetic polymorphism for thermal stability. *Arch Biochem Biophys* 1997;346:277–286. [PubMed: 9343375]
7. Salavaggione OE, et al. Thiopurine S-methyltransferase pharmacogenetics: variant allele functional and comparative genomics. *Pharmacogenet Genomics* 2005;15:801–815. [PubMed: 16220112]
8. Luka Z, et al. Mutations in human glycine N-methyltransferase give insights into its role in methionine metabolism. *Hum Genet* 2002;110:68–74. [PubMed: 11810299]
9. Luka Z, Wagner C. Effect of naturally occurring mutations in human glycine N-methyltransferase on activity and conformation. *Biochem Biophys Res Commun* 2003;312:1067–1072. [PubMed: 14651980]
10. Item CB, et al. Characterization of seven novel mutations in seven patients with GAMT deficiency. *Hum Mutat* 2004;23:524. [PubMed: 15108290]
11. Ji Y, et al. Human phenylethanolamine N-methyltransferase pharmacogenomics: gene re-sequencing and functional genomics. *J Neurochem* 2005;95:1766–1776. [PubMed: 16277617]
12. da Costa KA, et al. Common genetic polymorphisms affect the human requirement for the nutrient choline. *FASEB J* 2006;20:1336–1344. [PubMed: 16816108]
13. Wood TC, et al. Human arsenic methyltransferase (AS3MT) pharmacogenetics: gene resequencing and functional genomics studies. *J Biol Chem* 2006;281:7364–7373. [PubMed: 16407288]
14. Spielman RS, Weinshilboum RM. Genetics of red cell COMT activity: analysis of thermal stability and family data. *Am J Med Genet* 1981;10:279–290. [PubMed: 7304673]
15. Lachman HM, et al. Human catechol-O-methyltransferase pharmacogenetics: description of a functional polymorphism and its potential application to neuropsychiatric disorders. *Pharmacogenetics* 1996;6:243–250. [PubMed: 8807664]
16. Chen J, et al. Functional analysis of genetic variation in catechol-O-methyltransferase (COMT): effects on mRNA, protein, and enzyme activity in postmortem human brain. *Am J Hum Genet* 2004;75:807–821. [PubMed: 15457404]
17. Wedren S, et al. Catechol-O-methyltransferase gene polymorphism and post-menopausal breast cancer risk. *Carcinogenesis* 2003;24:681–687. [PubMed: 12727796]
18. Karayiorgou M, et al. Family-based association studies support a sexually dimorphic effect of COMT and MAOA on genetic susceptibility to obsessive-compulsive disorder. *Biol Psychiatry* 1999;45:1178–1189. [PubMed: 10331110]
19. Wang L, Weinshilboum R. Thiopurine S-methyltransferase pharmacogenetics: insights, challenges and future directions. *Oncogene* 2006;25:1629–1638. [PubMed: 16550163]
20. Horton JR, et al. Two polymorphic forms of human histamine methyltransferase: structural, thermal, and kinetic comparisons. *Structure* 2001;9:837–849. [PubMed: 11566133]

21. Rutherford K, et al. Crystal structures of human 108V and 108M catechol O-methyltransferase. *J Mol Biol* 2008;380:120–130. [PubMed: 18486144]
22. DeVry CG, Clarke S. Polymorphic forms of the protein L-isoaspartate (D-aspartate) O-methyltransferase involved in the repair of age-damaged proteins. *J Hum Genet* 1999;44:275–288. [PubMed: 10496068]
23. Tai HL, et al. 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]
24. Wang L, et al. Human thiopurine S-methyltransferase pharmacogenetics: variant allozyme misfolding and aggresome formation. *Proc Natl Acad Sci U S A* 2005;102:9394–9399. [PubMed: 15967990]
25. Rutherford K, et al. The V108M mutation decreases the structural stability of catechol O-methyltransferase. *Biochim Biophys Acta* 2008;1784:1098–1105. [PubMed: 18474266]
26. Rutherford K, et al. The histamine N-methyltransferase T105I polymorphism affects active site structure and dynamics. *Biochemistry* 2008;47:893–901. [PubMed: 18154359]
27. Rutherford K, Daggett V. A hotspot of inactivation: The A22S and V108M polymorphisms individually destabilize the active site structure of catechol O-methyltransferase. *Biochemistry* 2009;48:6450–6460. [PubMed: 19435324]
28. Doyle AE, Yager JD. Catechol-O-methyltransferase: effects of the val108met polymorphism on protein turnover in human cells. *Biochim Biophys Acta* 2008;1780:27–33. [PubMed: 17980711]
29. Rutherford K, Daggett V. The V119I polymorphism in protein L-isoaspartate O-methyltransferase alters the substrate-binding interface. *Protein Eng Des Sel* 2009;22:713–721. [PubMed: 19801578]
30. Tai HL, et al. 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]
31. Rutherford K, Daggett V. Four human thiopurine s-methyltransferase alleles severely affect protein structure and dynamics. *J Mol Biol* 2008;379:803–814. [PubMed: 18482735]
32. Wedren S, et al. Catechol-O-methyltransferase gene polymorphism and post-menopausal breast cancer risk. *Carcinogenesis* 2003;24:681–687. [PubMed: 12727796]
33. Karayiorgou M, et al. Family-based association studies support a sexually dimorphic effect of COMT and MAOA on genetic susceptibility to obsessive-compulsive disorder. *Biol Psychiatry* 1999;45:1178–1189. [PubMed: 10331110]
34. Meyer-Lindenberg A, et al. Midbrain dopamine and prefrontal function in humans: interaction and modulation by COMT genotype. *Nat Neurosci* 2005;8:594–596. [PubMed: 15821730]
35. Horton JR, et al. Structural basis for inhibition of histamine N-methyltransferase by diverse drugs. *J Mol Biol* 2005;353:334–344. [PubMed: 16168438]
36. Ogawa H, et al. Structure, function and physiological role of glycine N-methyltransferase. *Int J Biochem Cell Biol* 1998;30:13–26. [PubMed: 9597750]
37. Mudd SH, et al. Glycine N-methyltransferase deficiency: a novel inborn error causing persistent isolated hypermethioninaemia. *J Inherit Metab Dis* 2001;24:448–464. [PubMed: 11596649]
38. Augoustides-Savvopoulou P, et al. Glycine N-methyltransferase deficiency: a new patient with a novel mutation. *J Inherit Metab Dis* 2003;26:745–759. [PubMed: 14739680]
39. Luka Z, et al. Destabilization of human glycine N-methyltransferase by H176N mutation. *Protein Sci* 2007;16:1957–1964. [PubMed: 17660255]
40. Axelrod J. Purification and properties of phenylethanolamine-N-methyl transferase. *J Biol Chem* 1962;237:1657–1660. [PubMed: 13863458]
41. Bondinell WE, et al. Inhibitors of phenylethanolamine N-methyltransferase and epinephrine biosynthesis: a potential source of new drugs. *Drug Metab Rev* 1983;14:709–721. [PubMed: 6352222]
42. Saito S, Iida A, Sekine A, Miura Y, Sakamoto T, Ogawa C, Kawauchi S, Higuchi S, Nakamura Y. Identification of 197 genetic variations in six human methyltransferase genes in the Japanese population. *J Hum Genet* 2001;46:529–537. [PubMed: 11558902]

43. Cui J, et al. Association of polymorphisms in the promoter region of the PNMT gene with essential hypertension in African Americans but not in whites. *Am J Hypertens* 2003;16:859–863. [PubMed: 14553966]
44. Mann MB, et al. Phenylethanolamine N-methyltransferase (PNMT) gene and early-onset Alzheimer disease. *Am J Med Genet* 2001;105:312–316. [PubMed: 11378842]
45. Mann MB, et al. Association between the phenylethanolamine N-methyltransferase gene and multiple sclerosis. *J Neuroimmunol* 2002;124:101–105. [PubMed: 11958827]
46. Mudd SH, Poole JR. Labile methyl balances for normal humans on various dietary regimens. *Metabolism* 1975;24:721–735. [PubMed: 1128236]
47. Leuzzi V, et al. A mutation on exon 6 of guanidinoacetate methyltransferase (GAMT) gene supports a different function for isoform a and b of GAMT enzyme. *Mol Genet Metab* 2006;87:88–90. [PubMed: 16293431]
48. Hernandez A, et al. Role of the Met(287)Thr polymorphism in the AS3MT gene on the metabolic arsenic profile. *Mutat Res* 2008;637:80–92. [PubMed: 17850829]
49. Dong H, et al. The phosphatidylethanolamine N-methyltransferase gene V175M single nucleotide polymorphism confers the susceptibility to NASH in Japanese population. *J Hepatol* 2007;46:915–920. [PubMed: 17391797]
50. Pettersen EF, et al. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* 2004;25:1605–1612. [PubMed: 15264254]
51. Smith, CD., et al. Crystal structure of human L-isoaspartyl-O-methyl-transferase with S-adenosyl homocysteine at 1.6-Å resolution and modeling of an isoaspartyl-containing. 2002.
52. Wu H, et al. Structural basis of allele variation of human thiopurine-S-methyltransferase. *Proteins* 2007;67:198–208. [PubMed: 17243178]
53. Huang Y, et al. Mechanisms for auto-inhibition and forced product release in glycine N-methyltransferase: crystal structures of wild-type, mutant R175K and S-adenosylhomocysteine-bound R175K enzymes. *J Mol Biol* 2000;298:149–162. [PubMed: 10756111]
54. Luka Z, Wagner C. Human glycine N-methyltransferase is unfolded by urea through a compact monomer state. *Arch Biochem Biophys* 2003;420:153–160. [PubMed: 14622985]
55. Pakhomova S, et al. Glycine N-methyltransferases: a comparison of the crystal structures and kinetic properties of recombinant human, mouse and rat enzymes. *Proteins* 2004;57:331–337. [PubMed: 15340920]

Figure 1. Polymorphisms in Human SAM-Dependent Methyltransferases

(A) The conserved SAM-binding domain (blue) consists of a core 7-stranded β -sheet sandwiched between two sets of helices. The substrate specificity of each protein is determined by the non-conserved structural elements (orange). Several common coding polymorphisms that alter enzymatic activity and protein stability have been identified in COMT, HNMT, PIMT, TPMT, GNMT, GAMT, and PNMT. These variant residues are not localized to the active-site, and instead are distributed throughout the protein structure. SAM and SAH are colored in magenta and variant residues are colored in gold. (B) COMT, HNMT, PIMT, and TPMT all contain a common SNP-associated substitution site located 13–20 Å from the SAM-binding site at the intersection of $\alpha 2$, $\alpha 3$, and $\beta 3$. The variant residues are connected to the active site via $\alpha 3$ or $\beta 2$, which contain a conserved acidic residue (E90, E89, D109, E90, colored in green) that forms hydrogen bonds with the ribose hydroxyl group of SAM. Variant residues are shown in space-filling representation and colored in gold. SAM and SAH are colored in magenta.

Figure 2. Polymorphisms of COMT, HNMT, PIMT, and TPMT Distort the SAM- and Substrate-Binding Sites

The introduction of a larger residue in the polymorphic sites of COMT (V108M), HNMT (T105I), and PIMT (V119I) alters side-chain interactions between the variants residue and residues in α -2, α -3 and β -3. The resulting burial of the polymorphic residue within a hydrophobic pocket affects the orientation of α -3, α -4 and β -3, increasing the flexibility of several SAM-binding residues and loosening up the active site. These changes are translated throughout the protein, distorting the substrate-binding sites. The A80P substitution in TPMT disrupts the last turn of α 2, resulting in a large scale reorientation of α 2, coupled with increased fluctuations of α 1 and α 3 and the creation of a large cleft in the protein. Crystal structures of

COMT (3bwm.pdb21), HNMT (2aot.pdb35), PIMT (1i1n.pdb51), and TPMT (2bzg.pdb52) bound with SAM or SAH are shown in ribbon representation (left). Structures of the wild-type and variant proteins from the final ns of MD simulations at 37°C are shown in the center and right-hand columns, respectively. Variant residues are shown in space-filling representation and colored in yellow. SAM- and substrate-binding residues are shown in stick representation and colored green and red, respectively. Helices and strands reoriented by the variant residue are colored in blue. SAM and SAH are colored magenta.

Figure I. Structural Effects of the L49P, N140S, and H176N GNMT Polymorphisms

(A) Ribbon structure of GNMT showing the dimer and tetramer interfaces. Residues 140 and 176 are located in the active site, whereas residue 49 is found in α 1. Substitution of proline for leucine at position 49 could break helix 1, altering the orientation of α 1. Any rearrangement in the position of α 1 could both disrupt the dimer interface and move W30 and Y33 out of the active site. Helices A and 1 are colored magenta. Variant residues are shown in space-filling representation and colored gold. Citrate and residue Y33 are shown in stick representation and colored by atom type. (B) Superposition of active-site residues taken from crystal structures of the wild-type (green, 1R7455) and H176N (gold, 2AZT39) GNMT proteins. The active-site structures of the two proteins differ by 0.3 Å C α -rmsd, with residues H176N, H144, R177, and

H288 adopt slightly different conformations in the mutant protein. However, the overall active-site structures of the two proteins differ by only 0.3 Å C α -rmsd and the key active-site residues are able to bind citrate. Residue N140 forms van der Waals contacts with SAM and a hydrogen bond with the glycine substrate. The N140S substitution could disrupt these interactions and inactivate GNMT.