Computational Biology and Chemistry 59 (2015) 48–55

Contents lists available at ScienceDirect

Computational Biology and Chemistry

journal homepage: www.elsevier.com/locate/compbiolchem

Research article

Structural and functional impact of missense mutations in TPMT: An integrated computational approach

Esmat Fazel-Najafabadi[a], Elham Vahdat Ahar[b], Shirin Fattahpour[a], Maryam Sedghi[a,c,*]

[a] Medical Genetics Laboratory, Alzahra University Hospital, Isfahan University of Medical Sciences, Isfahan, Iran
[b] Institute of Biochemistry and Biophysics, University of Tehran, Iran
[c] Pediatric Inherited Diseases Research Center, Research Institute for Primordial Prevention of Non-communicable disease, Isfahan University of Medical Sciences, Isfahan, Iran

ARTICLE INFO

Article history:
Received 9 May 2015
Received in revised form 17 August 2015
Accepted 6 September 2015
Available online 9 September 2015

Keywords:
Bioinformatic analysis
TPMT
Missense mutations
Structure-function relationship

ABSTRACT

Background: Thiopurine S-methyltransferase (TPMT) detoxifies thiopurine drugs which are used for treatment of various diseases including inflammatory bowel disease (IBD), and hematological malignancies. Individual variation in TPMT activity results from mutations in *TPMT* gene. In this study, the effects of all the known missense mutations in TPMT enzyme were studied at the sequence and structural level

Methods: A broad set of bioinformatic tools was used to assess all the known missense mutations affecting enzyme activity. The effects of these mutations on protein stability, aggregation propensity, and residue interaction network were analyzed.

Results: Our results indicate that the missense mutations have diverse effects on TPMT structure and function. Stability and aggregation propensities are affected by various mutations. Several mutations also affect residues in ligand binding site.

Conclusions: In vitro study of missense mutation is laborious and time-consuming. However, computational methods can be used to obtain information about effects of missense mutations on protein structure. In this study, the effects of most of the mutations on enzyme activity could be explained by computational methods. Thus, the present approach can be used for understanding the protein structure-function relationships.

1. Introduction

TPMT is a cytoplasmic enzyme catalyzes the S-methylation of thiopurine drugs used for treatment of inflammatory bowel disease (IBD), hematological malignancies, and as immunosuppressant after organ transplantation. 6-Mercaptopurine, 6-thioguanine and azathiopurine are the drugs which their metabolism is affected by the low and intermediate TPMT enzyme activity (Wang et al., 2010).

TPMT gene is located on chromosome 6 (6p22.3) and contains 10 exons (Szumlanski et al., 1996). TPMT enzyme contains 8 helices and 9 sheets shown by letters (A–H) and number (1–9), respectively (Fig. 1) (Wu et al., 2007; Stivala et al., 2011). Genetic polymorphisms in TPMT enzyme result in decreased or absence of enzymatic activity in individuals carrying these alleles (Weinshilboum, 2001). To date, 43 alleles of TPMT designed from *1 to *37 are identified in human population (Appell et al., 2013). TPMT*1 is the wild type form of the enzyme which has the normal activity. Reduced enzyme activity was observed in individuals carrying a normal allele and a defective allele (Relling et al., 2013).

Despite the introduction of TPMT as a drug metabolizer for approximately half a century, there is no information about the natural substrate of this enzyme. To date, four substrates of TPMT in the metabolic pathway of thiopurine drugs are identified (Zaza et al., 2010). 6-Mercaptopurine (6MP), 6-thioguanine (6-TG), thioguanine monophosphate (6-TGMP) and thioinosine monophosphate (6TIMP) are identified as methyl acceptor and S-adenosyl-L-methionine (SAM) as methyl donor (Ujiie et al., 2008; Peng et al., 2008). Considerable progress has been made in understanding the effects of several TPMT alleles on protein activity using heterologous expression systems (Ujiie et al., 2008; Garat et al., 2008; Salavaggione et al., 2005; Hamdan-Khalil et al., 2005). However, there are only a few reports illustrating the effects of most common mutations on TPMT enzyme structure using in vitro and computational approaches (Wu et al., 2007; Garat et al.,

* Corresponding author at: Medical Genetics Laboratory, Alzahra University Hospital, Isfahan University of Medical Sciences, Isfahan, Iran. Fax: +98 3136201281.
E-mail address: m_sedghi@azh.mui.ac.ir (M. Sedghi).

http://dx.doi.org/10.1016/j.compbiolchem.2015.09.004
1476-9271/

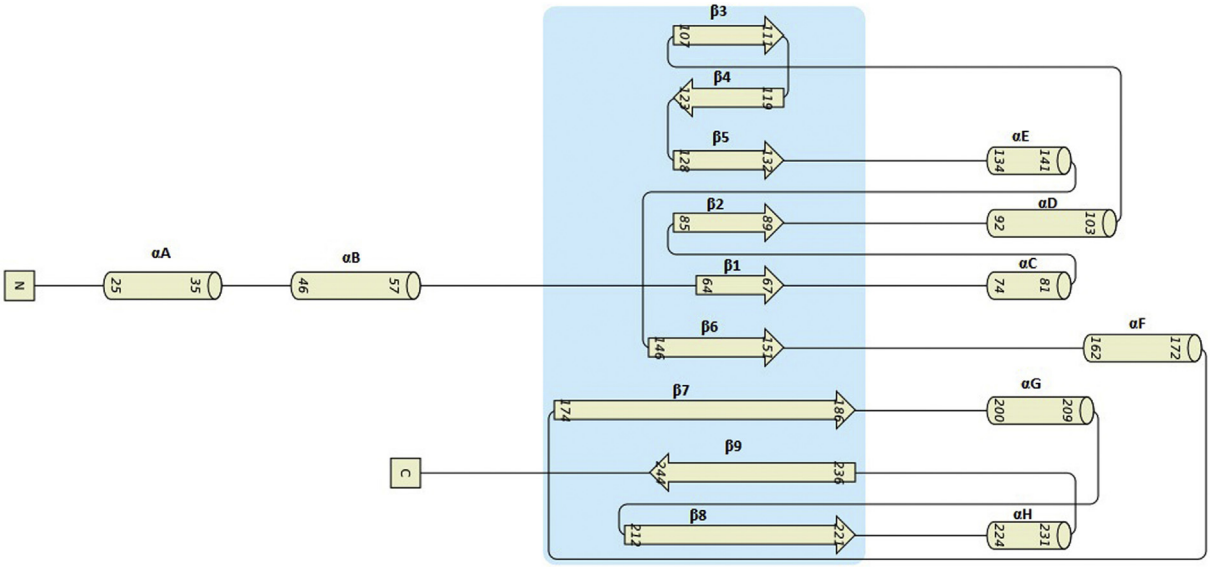

Fig. 1. Schematic representation of TPMT enzyme. Residue numbers indicate the range of each secondary structure. Cylinders are helices (letters from A–H) and broad arrows (numbers from 1 to 9) are strands (image was obtained from http://munk.csse.unimelb.edu.au/pro-origami/(Stivala et al., 2011) using 2 h11 chain A as input).

2008; Wennerstrand et al., 2012; Rutherford and Daggett, 2008). Bioinformatic analysis of point mutations is used routinely by researchers and clinicians to advise the pathogenecity of mutations in combination with other evidence in a variety of diseases (Heineman et al., 2015; Nouri et al., 2014; Keivani et al., 2015). There are different approaches used by researchers, but the most reliable results can be obtained by applying an extensive set of prediction programs. Most of the programs accept both sequence and structure as input; however, using structure information when one is available yields a more complete prediction. For example, structural and functional effects of missense mutations in GJB2 gene was predicted by several in silico tools using both sequence and structure information of the human Connexin 26. Researchers can use the results obtained by this study in their in vitro studies to provide further insights into molecular mechanisms of deafness (Yilmaz, 2015). In the present study, we used bioinformatic tools to study the effects of TPMT variants on protein structure. With the exception of M1V and M1T which result in lack of TPMT translation, all the missense mutation identified in TPMT were modeled onto the corresponding structure and their effects on protein structure were evaluated. In addition, sequence-based analysis of TPMT was used for illustration of the effects of TPMT variants on protein aggregation.

2. Methods

2.1. Collection of missense mutations and protein information

Information about TPMT alleles was obtained from http://www.imh.liu.se/tpmtalleles/tabell-over-tpmt-alleler?l=en (Appell et al., 2013). The effects of mutations on enzyme activity were retrieved from the PharmGKB (Whirl-Carrillo et al., 2012) and the literature (Ujiie et al., 2008; Salavaggione et al., 2005).

2.2. Amino acid conservation

Polyview (http://polyview.cchmc.org/) (Porollo et al., 2004) was used for scoring conserved residues in TPMT protein. Polyview is a web server program for annotation and 2D visualization of protein structures which represent information such as conserved residues and residue relative solvent accessibility. For multiple sequence alignment visualization, sequence homologues to human TPMT protein were obtained from NCBI http://www.ncbi.nlm.nih.gov/ Multiple sequence alignment was performed by Clustal W omega (http://www.ebi.ac.uk/Tools/msa/clustalo/) (Sievers et al., 2011) and then visualized by Jalview (Waterhouse et al., 2009). Jalview is available as a Java applet for on-line use and a desktop application for visualization and annotation of multiple sequence alignments.

2.3. Effects of nsSNPs on protein aggregation and stability

AMYLPRED2 (http://aias.biol.uoa.gr/AMYLPRED2/) (Tsolis et al., 2013) was used to detect the effects of mutation on TPMT aggregation. AMYLPRED2 integrates 11 algorithms for prediction of aggregation prone regions in proteins. Effect of mutations on chaperone binding was investigated using LIMBO (http://limbo.switchlab.org/limbo-analysis) (Van Durme et al., 2009). LIMBO algorithm is developed based on information from peptide binding experiments using the prokaryotic heat-shock protein DnaK which is a representative of the ubiquitous Hsp70 family. These molecular chaperones are specialized in the binding to exposed hydrophobic regions in unfolded polypeptides which result in protein quality control. Stability changes of protein upon amino acid substitutions were predicted by four programs including, SDM (http://mordred.bioc.cam.ac.uk/~sdm/sdm.php) (Worth et al., 2011), I-Mutant2.0 (http://folding.biofold.org/i-mutant/i-mutant2.0.html) (Capriotti et al., 2005) I-Mutant3.0 (http://gpcr2.biocomp.unibo.it/cgi/predictors/I-Mutant3.0/I-Mutant3.0.cgi) (Capriotti et al., 2008) and FoldX (Guerois et al., 2002). FoldX predictions were obtained from SNPeffect 4.0 (http://snpeffect.switchlab.org/), an online predition tool using several bioinformatic tools to predict the effects of coding SNPs on protein structure (De Baets et al., 2012). For all these programs protein structure was used as input and default parameters were not changed.

2.4. Structural analysis

Effects of mutations on protein structure were studied by determination of clashing sides and Residueinteraction network (RIN) changes upon mutations. Crystal structure of the TPMT (PDB 2h11) was used for structural analysis. Mutated structures were

created and visualized by UCSF Chimera (Pettersen et al., 2004). The best rotamers were used for analysis. Clashes in modeled structures were determined using UCSF Chimera. Residue interaction network data for 2h11 structure was obtained from http://rinalyzer.de/rindata.php Comparison of residue interaction network of wildtype and mutated structures was visualized by Cytoscape (http://www.cytoscape.org/) (Saito et al., 2012). StructureViz (Morris et al., 2007) and RINalyzer (Doncheva et al., 2011) are two plugins of Cytoscape which links Cytoscape to UCSF Chimera and were used for RIN analysis.

TPMT is co-crystallized with the S-adenosyl L-homocystein (SAH) in PDB structure 2h11. We used molecular docking to identify residues in ligand binding site of TPMT. AmongTPMT substrates, 3D structures of SAM and 6MP were available at ligand Expo (http://ligand-expo.rcsb.org/). AUTODOCK Vina was used to mimic the binding of SAM, and 6MP into 2h11 PDB structure and to extract information about residues contacting with these ligands.

3. Results

Results obtained from prediction tools obtained in this study are summarized in Table 1. Conserved positions in a protein are essential for protein structure and function. Mutations in conserved position are likely to result in deleterious effects. TPMT is a highly conserved protein among mammals. Residues in positions A80, Y240, L49, E28, R163, G71, L69, F208, G36, I204, and F67 have a high score of conservation according to Polyview (Fig. 2A). Among other residues, eight of them including Y107, G117, K119, G144, A154, Y180, V199, and C212 are also conserved in alignment of homologous sequences obtained from NCBI (Fig. 2B). Among the mutations located in conserved positions, 16 are in buried regions of protein (L49S, L69V, F67S, G71R, A80P, A154T, Y180F, V199I, I204T, F208L, C212R, Y240C, R163H/P/C, G117R). In positions 69, 80, 199, and 208, 240 the hydrophobicity nature of residues are not changed after substitution because hydrophobic residues are introduced. R163 is substituted with three different amino acids, two non polar (Cys and Pro) and one with charged side chain (His). Each of these substituted amino acids has different physicochemical properties which might have different consequences after substitution (Fig. 3).

AMYLPRED2 reports the aggregation-prone regions when at least 5 out of 11 algorithms agree on the prediction. According to AMYLPRED2 L69, Y180, R215, C132, Q179, C212, F208, A154, Y240 are located in regions essential for aggregation. According to LIMBO prediction, residues in positions 204,208, 227, 238 and 240 are located in regions necessary for DnaK chaperon protein

Table 1
Summary of the effects of missense mutations in *TPMT* gene on enzyme structure and function.

AA	Allele	Highly conserved	Functional[a]	RIN[b]	Enzyme activity[c]		Clash
					Defective	Decreased	
A80P	TPMT*2	*		*	*(Ujiie et al., 2008)	*(Salavaggione et al., 2005)	*
Y240C, A154T	TPMT*3A		×	*	*(Ujiie et al., 2008; Salavaggione et al., 2005)		*
A154T	TPMT*3B		×	*	*(Ujiie et al., 2008; Salavaggione et al., 2005)		*
Y240C	TPMT*3C	*		*		*(Ujiie et al., 2008; Salavaggione et al., 2005)	
L49S	TPMT*5	*		*	*(Ujiie et al., 2008; Salavaggione et al., 2005)		*
Y180F	TPMT*6			*		*(Ujiie et al., 2008; Salavaggione et al., 2005)	*
H227Q	TPMT*7				*(Ujiie et al., 2008; Hamdan-Khalil et al., 2003)	*(Salavaggione et al., 2005)	*
R215H	TPMT*8			*		*(Ujiie et al., 2008; Salavaggione et al., 2005)	*
K119T	TPMT*9			*		*(Ujiie et al., 2008; Salavaggione et al., 2005)	
G144R	TPMT*10					*(Ujiie et al., 2008; Salavaggione et al., 2005; Hamdan-Khalil et al., 2003)	
C132Y	TPMT*11		×	*	*(Ujiie et al., 2008)	*(Salavaggione et al., 2005)	*
S125L	TPMT*12			*		*(Ujiie et al., 2008; Salavaggione et al., 2005; Hamdan-Khalil et al., 2003)	
E28V	TPMT*13	*	×	*		*(Ujiie et al., 2008; Salavaggione et al., 2005; Hamdan-Khalil et al., 2003)	
R163H	TPMT*16	*		*		*(Ujiie et al., 2008)	*
Q42E	TPMT*17		*	*	*(Ujiie et al., 2008)		*
G71R	TPMT*18	*	*		*(Ujiie et al., 2008)		*
K122T	TPMT*19			*		*(Ujiie et al., 2008)	
K238E	TPMT*20			*		*(Ujiie et al., 2008)	
L69V	TPMT*21	*	*	*	*(Ujiie et al., 2008; Garat et al., 2008)		*
R163P	TPMT*22	*		*	*(Ujiie et al., 2008)		*
A167G	TPMT*23			*		*(Ujiie et al., 2008)	
Q179H	TPMT*24			*		*(Garat et al., 2008)	
C212R	TPMT*25			*	*(Garat et al., 2008)		*
F208L	TPMT*26	*					*
Y107D	TPMT*27			*		*(Feng et al., 2010)	
G117R	TPMT*28						
G36S	TPMT*30	*	×	*	*(Ujiie et al., 2008)		
I204T	TPMT*31	*		*		*(Appell et al., 2010)	
E114K	TPMT*32			*			*
R163C	TPMT*33	*		*			
R82W	TPMT*34			*			*
F67S	TPMT*35	*	×	*			
V199I	TPMT*36			*			*

[a] Amino acids in direct contact with ligands (* residues contact with ligands directly, × residues contact with ligands indirectly).
[b] Residue interaction network changes after substitution.
[c] Enzyme activity of TPMT alleles observed in in vitro relative to wild type. Defective (0–10%), decreased more than 10%).

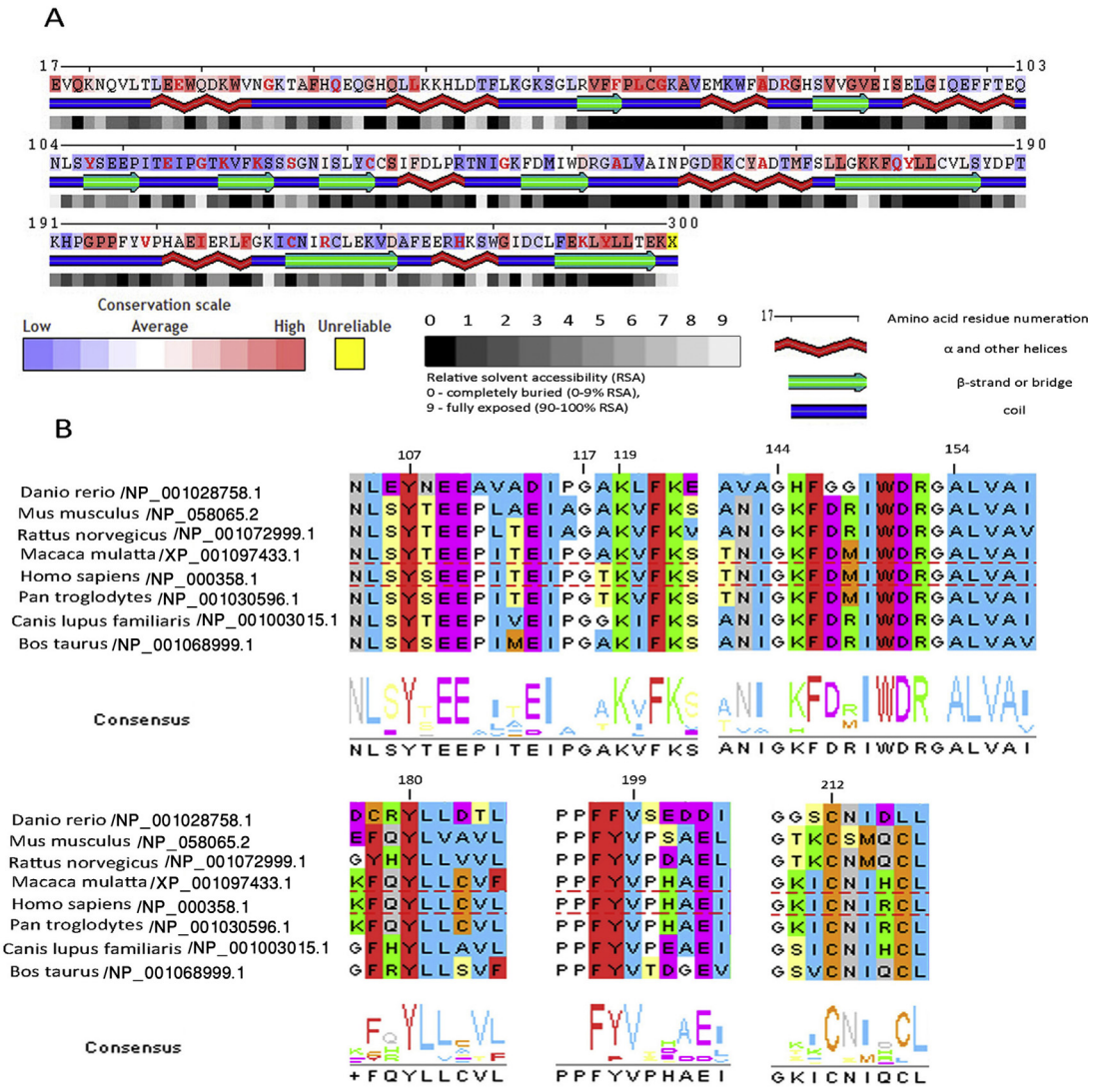

Fig. 2. Conserved residues in TPMT enzyme. (A) Visualizing the TPMT secondary structure elements, conserved residues, relative solvent accessibility using polyview (http://polyview.cchmc.org/). The mutated residues are shown in red. (B) TPMT multiple sequence alignment using clustal omega (http://www.ebi.ac.uk/Tools/msa/clustalo/). In this figure residues with low score indicated by polyview which were conserved in alignment of the human TPMT amino acid sequence with those of vertebrate orthologs are shown. Residues with similar physicochemical properties have the same color.

binding. Among these residues, I204T and Y240C cause TPMT to lose one and two DnaK chaperon protein binding stretches, respectively, which might decrease chaperone binding affinity of the protein. I204 is the second residue in EIERLFG region, and Y240 is in two overlapping stretches KLYLLTE and LYLLTEK.

SDM, I-Mutant2.0, FoldX and I-Mutant3.0 were used for prediction of changes in protein stability because of the mutations. FoldX predicted five and sixteen substitutions result in severe and intermediate stability decrease, respectively. According to I-Mutant2.0, 23 mutations result in decrease in protein stability, with nine mutations have intermediate effect and four one have severe effect. Large decrease in protein stability was predicted for 21 mutations by I-Mutant3.0. Using SDM prediction, twenty mutations were predicted to result in stability decrease of the TPMT in which 11 were reported to be highly destabilizing and disease causing. All the four prediction tools also report increase in protein stability after mutation. Among the programs used in this study, SDM and I-Mutant3.0 predicted that R163H and E28V highly stabilize protein, respectively. Thirteen mutations A80P, Y240C, L49S, R163H, L69V, R163P, A167G, C212R, F208L, Y107D, G36S, I204T, and F67S were predicted to have an intermediate/severe effect on protein stability by three programs (Table 2).

Clashing side chains and RIN change were used to study the effects of mutations on protein structure. A80P, L49S, Y180F, H227Q, R215H, C132Y, R163H, Q42E, G71R, L69V, R163P, C212R, F208L, E114K, R82W, V199I, and A154T do not fit into TPMT without causing serious clashing with other side chains in neighboring amino acids. Among these mutation F208L, Q42E, and V199I might have less severe effects on protein scaffolding. These mutated amino acids can be fit into TPMT after a few steps of structure minimization. H227Q, G144R, G71R, F208L and G117R are the only mutations that do not have significant effect on RIN.

To determine binding site of SAM, and 6MP, autodock VINA was used for docking. The experimentally determined structures and previous models were also used for identification of amino acids involved in ligand binding. Ala154, Cys132, Glu28, Gln42, Gly71, Leu69, Gly36, and Phe67 were involved directly or indirectly with modeled ligands. Residues with indirect contacts with ligands were specified as the ones having contacts with residues with direct contact with the ligands. The substitution of one of these

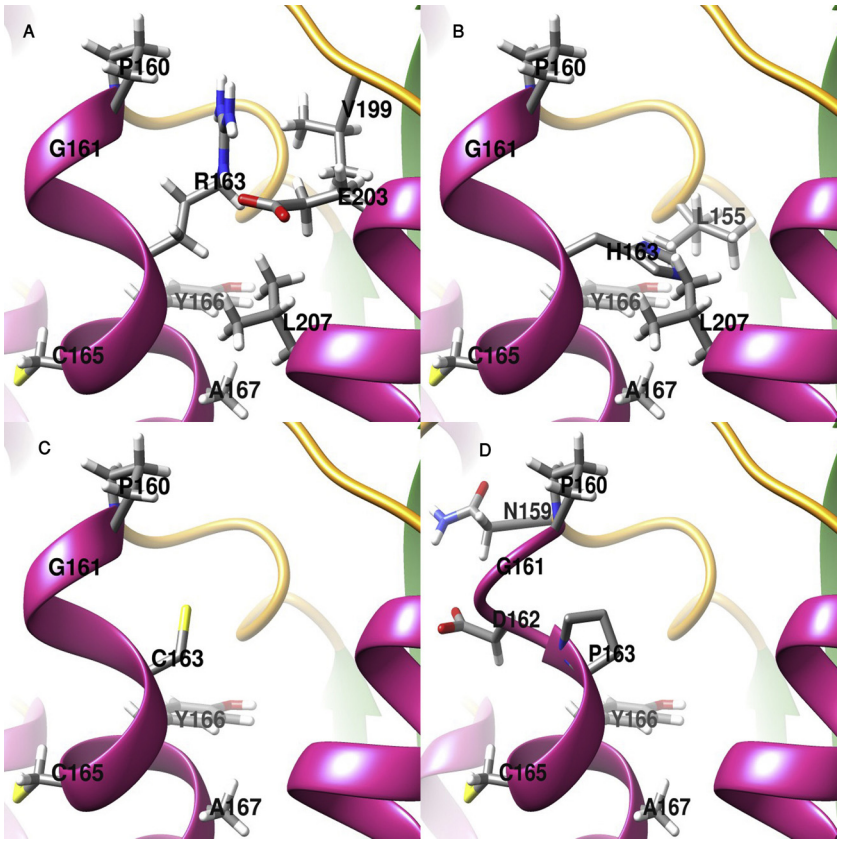

Fig. 3. Residue 163 is substituted with three different residues. Arg is the wild type residue in position 163. Three alleles of TPMT (TPMT*16, TPMT*22, TPMT*33) are resulted from mutations in position 163. (A) Arg 163 has contacts with eight residues including P160, G161, C165, Y166, A167, V199, E203, L207. (B) His in position 163 has contacts with P160, G161, C165, Y166, A167, L207, and L155. This position loses contact with V199 in the loop between β7 and αG. And has new contacts with L155 in the loop between β6 and αF. (C) After substitution of R163 with Cys, the amino acids in contact with this position are reduced to five including P160, G161, C165, Y166, and A167. (D) Replacement of Arg 163 with Pro 163 changes the secondary structure of helix.

amino acids is likely to have an effect on the specificity of the enzyme (Figs. 4 and 5). For example, Gln42 is in contact with Phe40 and Arg152, which both contact SAM indirectly. After mutating Gln to Glu, the contact of the position with Arg152 is lost. Alleles with this mutations had no enzyme activity in COS-7 expression system (Ujiie et al., 2008).

4. Discussion

Genetic polymorphisms in Human TPMT enzyme activity is now demonstrated to be correlated to interindividual differences in response to thiopurines. There are few studies investigating effects of substitutions on enzyme activity using in vitro systems (Ujiie et al., 2008; Garat et al., 2008; Salavaggione et al., 2005; Hamdan-Khalil et al., 2003). In addition, in vitro analysis of structural effects of TPMT mutations is only available for TPMT*3A, TPMT*3B, TPMT*3C, TPMT*2, and TPMT*5 (Wu et al., 2007; Wennerstrand et al., 2012). Assessment of the effects of TPMT variants on protein structure using computational methods are limited to a few alleles including TPMT*3A, TPMT*3B, TPMT*3C, TPMT*2, TPMT*9, TPMT*21, TPMT*23, TPMT*24, and TPMT*25 (Wu et al., 2007; Garat et al., 2008). In this study, the structural and functional effects of putative mutations in TPMT were investigated.

Residues in conserved positions are considered to be important for protein structure and function. And substituting these residues especially with residues with different physicochemical properties might lead to deleterious effects. In vitro studies showed that six substitutions A80P, L49S, G71R, L69V, R163P and G36S result in no detectable or significantly reduced enzyme activity (Ujiie et al., 2008; Salavaggione et al., 2005). These six substitutions are among 11 highly conserved residues identified in this study. For the remaining highly conserved residues including Y240C, E28V, R163H, F208L, I204T, R163C and F67S, decreased enzyme activity in in vitro expression systems was observed for Y240C, E28V, I204T and R163H (Ujiie et al., 2008; Salavaggione et al., 2005; Appell et al., 2010). Arginine in position 163 is substituted by three different residues Pro, His and Cys. R163P was reported to have no detectable enzyme activity in COS-7 and R163H had 16% of Vmax/Km of that wild type (Ujiie et al., 2008). Proline is a putative residue which changes the secondary structure of the protein after substitution; therefore, it is obvious that R163P result in enzyme instability. In R163H the physicochemical properties of the residue is not changed after substitution; however, the substituted residue has several clashes with neighboring residues and,consequently, lose contact with Val199 in the loop between β7 and αG which explain severe effect of this mutation on enzyme activity.

The effect of missense mutations on TPMT enzyme stability predicted by four programs is indicated in Table 2. Thirteen of the thirty two mutations affect the stability of protein structure. Seven of these mutations including A80P, Y240C, L49S, L69 V, R163P, C212R and G36S were reported to significantly decrease enzyme activity (0–10%). Experimental data indicates that the low enzyme activity of TPMT*5 is the result of local changes in the active site region rather than decreased stability (Wennerstrand et al., 2012). Even though R163H was predicted to reduce stability by three programs, SDM predicted this mutation highly increase protein stability. Generally, mutations decreasing protein stability are deleterious. Increased stability can reduce protein flexibility and

Table 2
Protein stability changes upon mutations.

Allele	SDM	I-Mutant 2.0	FoldX	I-Mutant 3.0
TPMT*2	↓↓↓	↓↓↓	↓↓↓	↓↓↓
TPMT*3B	↓	↓	↓↓↓	↓↓↓
TPMT*3C	↑	↓↓	↓↓	↓↓↓
TPMT*5	↓↓↓	↓↓↓	↓↓	↓↓↓
TPMT*6	↑↑	↓	↓	↓↓↓
TPMT*7	↓↓	↓	↓	N
TPMT*8	N	↓	↓	↓↓↓
TPMT*9	N	↓	↓↓	↓↓↓
TPMT*10	↓↓	N	↓↓	N
TPMT*11	↓↓↓	↓	↓↓	N
TPMT*12	N	↓↓	N	N
TPMT*13	N	↑	↓↓	↑↑↑
TPMT*16	↑↑↑	↓↓	↓↓	↓↓↓
TPMT*17	N	N	N	N
TPMT*18	↓	↓	↓↓	N
TPMT*19	N	↓	↓	↓↓↓
TPMT*20	↑↑	N	↓	N
TPMT*21	↓	↓↓	↓↓	↓↓↓
TPMT*22	↓↓	↓↓	↓↓↓	↓↓↓
TPMT*23	↓↓↓	↓↓↓	↓↓	↓↓↓
TPMT*24	↓	↓↓	N	↓↓↓
TPMT*25	↓↓↓	N	↓↓↓	↓↓↓
TPMT*26	N	↓↓	↓↓	↓↓↓
TPMT*27	↓↓↓	↓	↓↓	↓↓↓
TPMT*28	↓↓	N	↓↓	N
TPMT*30	↓↓↓	↓↓	↓↓	N
TPMT*31	↓↓↓	↓↓↓	↓↓	↓↓↓
TPMT*32	↓	↓↓	N	↓↓↓
TPMT*33	↑↑	N	↓↓	↓↓↓
TPMT*34	↓↓↓	N	N	N
TPMT*35	↓↓↓	↓↓↓	↓↓↓	↓↓↓
TPMT*36	↓	N	N	↓↓↓

↓↓↓ Severe decrease; ↓↓ intermediate decrease; ↓ mild decrease; ↑↑↑ severe increase; ↑↑ intermediate increase; ↑ mild increase; N neutral.
We used the following DDG range for describing the severity degree of stability changes obtained from Imutant2.0 (the output from the other programs includes a description for the severity of stability changes). DDG ≤ −2 kcal/mol → severe decrease; −1 kcal/mol ≥ DDG ≥ −2 kcal/mol → intermediate decrease; −1 kcal/mol > DDG ≥ −0.5 kcal/mol → mild decrease; −0.5 kcal/mol ≤ DDG ≤ 0.5 kcal/mol → Neutral; 0.5 kcal/mol ≤ DDG ≤ 1 kcal/mol → mild increase; 1 kcal/mol ≤ DDG ≤ 2 kcal/mol → intermediate increase; DDG > 2 kcal/mol → severe increase.

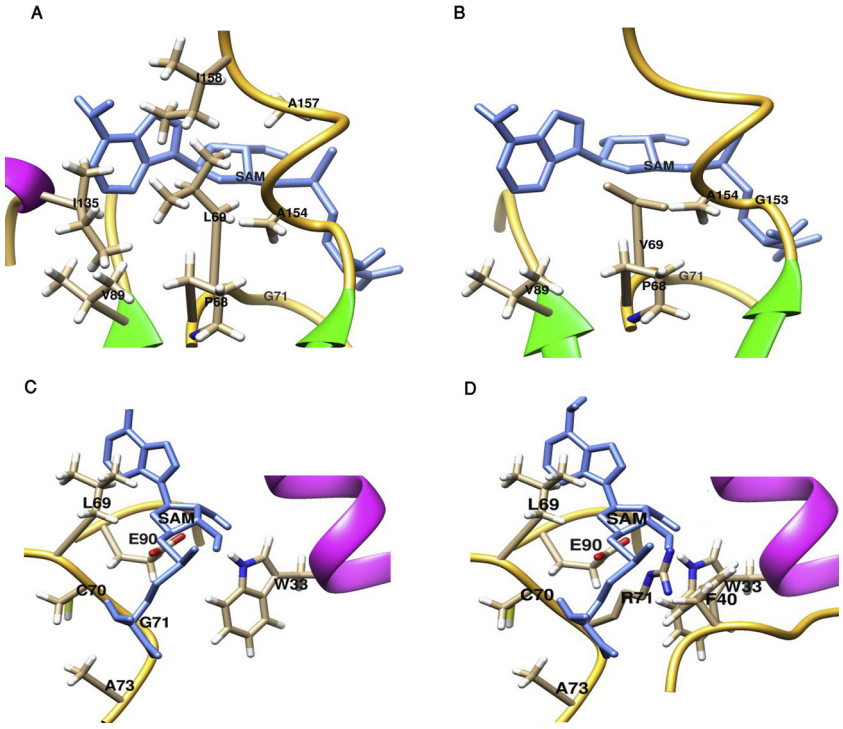

Fig. 4. Substituting residues in ligand binding site (L69 V and G71R). (A) L69 located in the loop between β1 and αC interacts with A154, A157 and I158 located in the loop between β6 and αF. (B) After mutating Leu in position 69 to Val, it loses interactions with A157 and I158 and also I135 in αE helix. The loss of these interactions might destabilize protein. (C) Gly is the wild type residue in position 71. G71 is in the same loop as L69 and interacts with C70, A73, L69, E90 and W33. (D) Substituting Gly to Arg in position 71. R71 acts like a barrier between F40 and its ligand. Phenylalanine in position 40 is located in the loop between αA and αB and is important in ligand binding.

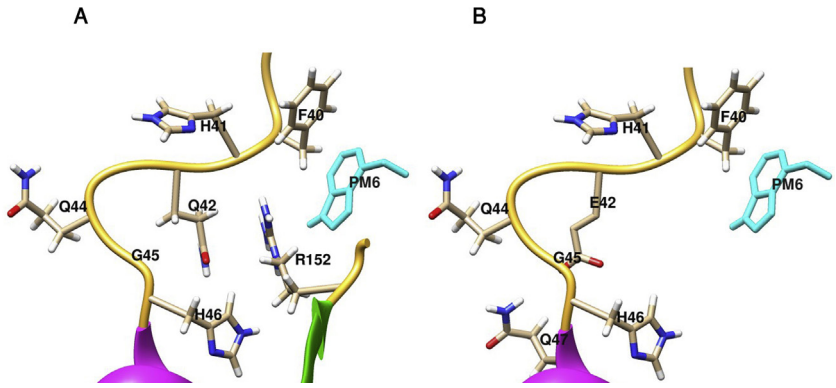

Fig. 5. Replacement of Gln with Glu in position 42. (A) In wildtype TPMT, Q42 interacts with 6 residues located in the loop between between αA and αB and the loop between β6 and αF. (B) In TPMT*17, Q42 has clashes with G45 which are missed after a few steps of structure minimizations; however, position 42 loses contacts with R152 in the loop between β6 and αF.

enzymatic activity in some conditions(Khan and Vihinen, 2010). E28V was also predicted to severely increase protein stability by Imutant3.0. E28 residue is in contact with ligand indirectly.

In previous studies it was suggested that many TPMT variants including wild type TPMT has a tendency to aggregate (Feng et al., 2010). In vitro studies for TPMT*2, TPMT*3A, TPMT*3B, TPMT*3C and TPMT*27 indicated a higher tendency of these alleles than wild type to aggregate in cultured cells (Wennerstrand et al., 2012; Feng et al., 2010; Wang et al., 2005). Aggresome formation in COS-1 cells was higher for TPMT*3A including both substitutions, A154T and Y240C, than TPMT*3B and TPMT*3C, respectively (Wang et al., 2005). Wennerstrand et al. (2012) showed the strong tendency of TPMT*2 to aggregate by studying biophysical properties of this variant. Degradation of TPMT*3A was shown to be increased by three pathways including a chaperone protein-dependent, proteasome-mediated pathway (Wang et al., 2005; Tai et al., 1997; Wang et al., 2003) and autophagy (Li et al., 2008). TPMT*27 is another TPMT allozyme which was shown to be degraded by both proteasome-mediated degradation and autophagy (Feng et al., 2010). In the present study, Tyr240 and Ala154 were among amino acids predicted to be located in aggregation prone stretches in TPMT; however, Tyr107 and Ala80 were not among them. According to LIMBO prediction, there are eight stretches in TPMTwt that bind to chaperones. Among the residues predicted to be located in these stretches, mutating Tyr240 and Ile204 leads to loss of two stretches, separately. So far, experimental studies showed that one of the mechanisms leading to increase degradation of TPMT*3A is chaperone dependent (Wang et al., 2003).

Residues with larger side chains than the original ones are likely to result in structural rearrangements. Side chains which cannot be fitted into protein without causing clashes lead to local rearrangements of the structure which may have an influence on protein stability. Rotamer analysis was performed to determine the clashes caused by the substituted amino acids. Amino acids which fitted into the structure without causing clashes were further analyzed to detect their effect on residue interaction network. Among mutations causing clashes, three of them including Y180F, C132Y and G71R have severe clashes with the side chains of neighboring amino acids. Ujiie et al. (2008) reported Y180F and C132Y have 17% and 3% of the Vmax/Km value for wild-type TPMT. G71R was also reported to have no detectable enzyme activity in S-methylation of 6TG. Further analysis of G71R clashing side chains indicated that these clashes are between the substituted Arg and ligands (6MP, and SAM) and a residue (phe40) located in the loop between αA and αB. This flexible loop previously was shown that is over the active site (the active site loop) and change conformation upon 6MP binding using crystal structure from murine TPMT (PDB 3BGI & 3BGD) (Peng et al., 2008). Based on sequence homology, the active site loop (Ser34-His47) in murine TPMT corresponds to Ala39-His52 in the human TPMT. The flexibility of this loop might be affected by replacement of Arg in position 71.

Three mutations Ala154, Cys132, Glu28, Gln42, Gly71, Leu69, Gly36, and Phe67 can be classified as functional mutations, with L69V, G71R directly changing amino acids involved in ligand –protein interactions. All these residues are highly conserved except for Ala154, Cys132, and Gln42. Replacement of Gln with Glu had no effect on stability; however, it results in clashes between Glu and Gly45 in the active site loop (between αA and αB). The other residue, Leu69, interacts with SAM via Hydrogen bond. Leu69 interacts with 3 other residues interacting with SAM and 6MP including Gly71, Ile135, and Ala 157. The side chains of substituted residue (Val) in position 69 have clashes with Ala154 and Pro68. These clashes may result in local rearrangement of the loop between β6 and αF containing Arg152, Gly153, and Ala157 which are conserved residues interacting with SAM. L69 V had no detectable enzyme activity in COS-7 cells in the presence of 6TG (Ujiie et al., 2008).

5. Conclusions

In this study we showed that most of the missense mutations resulting in low enzyme activity are located in strictly conserved positions. Mutations causing severe clashes have less enzyme activities in contrast to the mutations with few clashes. In addition, changes in residue contacts could be used to illustrate the effects of mutations. The data of enzyme activity in in in vitro is not available for a few missense mutations in TPMT; therefore, our results obtained from computational methods cannot be verified with experimental evidence. G117R and G144R were two only substitutions that their deleterious effects on TPMT enzyme could be defined by none of the prediction tools used in this study. Clearly, more biochemical assays are required to understand the effect of the remaining TPMT variants on enzyme activity. In addition, computational methods, especially prediction of aggregation prone regions, need to be improved so that can be used to precisely illustrate the TPMT structure-function relationship.

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgment

This work was funded by grant number 293047 from deputy for research, Isfahan University of Medical Sciences, Isfahan, Iran.

References

Appell, M., et al., 2010. Characterization of a novel sequence variant, TPMT*28, in the human thiopurine methyltransferase gene. Pharmacogenet. Genomics 20, 700–707.

Appell, M.L., et al., 2013. Nomenclature for alleles of the thiopurine methyltransferase gene. Pharmacogenet. Genomics 23, 242–248.

Capriotti, E., Fariselli, P., Casadio, R., 2005. I-Mutant2. 0: predicting stability changes upon mutation from the protein sequence or structure. Nucleic Acids Res. 33, W306–W310.

Capriotti, E., et al., 2008. A three-state prediction of single point mutations on protein stability changes. BMC Bioinformatics 9, S6.

De Baets, G., et al., 2012. SNPeffect 4.0: on-line prediction of molecular and structural effects of protein-coding variants. Nucleic Acids Res. 40, D935–D939.

Doncheva, N., et al., 2011. Analyzing and visualizing residue networks of protein structures. Trends Biochem Sci 36, 179–182.

Feng, Q., et al., 2010. Thiopurine S-methyltransferase pharmacogenetics: functional characterization of a novel rapidly degraded variant allozyme. Biochem. Pharmacol. 79 p. 1053.

Garat, A., et al., 2008. Characterisation of novel defective thiopurine S-methyltransferase allelic variants. Biochem. Pharmacol. 76, 404–415.

Guerois, R., Nielsen, J.E., Serrano, L., 2002. Predicting changes in the stability of proteins and protein complexes: a study of more than 1000 mutations. J. Mol. Biol. 320, 369–387.

Hamdan-Khalil, R., et al., 2003. In vitro characterization of four novel non-functional variants of the thiopurine S-methyltransferase. Biochem. Biophys. Res. Commun. 309, 1005–1010.

Hamdan-Khalil, R., et al., 2005. Identification and functional analysis of two rare allelic variants of the thiopurine S-methyltransferase gene, TPMT*16 and TPMT*19. Biochem. Pharmacol. 69, 525–529.

Heineman, T., et al., 2015. In silico analysis of NF2 gene missense mutations in neurofibromatosis type 2: from genotype to phenotype. Otol. Neurotol. 36, 908–914.

Keivani, A., et al., 2015. A new compound heterozygous mutation in GJB2 causes nonsyndromic hearing loss in a consanguineous Iranian family. Int. J. Pediatr. Otorhinolaryngol. 79, 553–556.

Khan, S., Vihinen, M., 2010. Performance of protein stability predictors. Hum Mutat 31, 675–684.

Li, F., et al., 2008. Thiopurine S-methyltransferase pharmacogenetics: autophagy as a mechanism for variant allozyme degradation. Pharmacogenet. Genomics. 18, 1083–1094.

Morris, J., et al., 2007. structureViz: linking cytoscape and UCSF chimera. Bioinformatics 23, 2345–2347.

Nouri, N., et al., 2014. Use of in silico tools for classification of novel missense mutations identified in dystrophin gene in developing countries. Gene 535, 250–254.

Peng, Y., et al., 2008. Structural basis of substrate recognition in thiopurine S-methyltransferase. Biochemistry 47, 6216–6225.

Pettersen, E., et al., 2004. UCSF Chimera—a visualization system for exploratory research and analysis. J. Comput. Chem. 25, 1605–1612.

Porollo, A., Adamczak, R., Meller, J., 2004. POLYVIEW: a flexible visualization tool for structural and functional annotations of proteins. Bioinformatics 20, 2460–2462.

Relling, M., et al., 2013. Clinical pharmacogenetics implementation consortium guidelines for thiopurine methyltransferase genotype and thiopurine dosing: 2013 update. Clin Pharmacol Ther 93, 324–325.

Rutherford, K., Daggett, V., 2008. Four human thiopurine s-methyltransferase alleles severely affect protein structure and dynamics. J. Mol. Biol. 379, 803–814.

Saito, R., et al., 2012. A travel guide to cytoscape plugins. Nat. Methods .

Salavaggione, O., et al., 2005. Thiopurine S-methyltransferase pharmacogenetics: variant allele functional and comparative genomics. Pharmacogenet. Genomics 15, 801–815.

Sievers, F., et al., 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol. 7, 539.

Stivala, A., et al., 2011. Automatic generation of protein structure cartoons with Pro-origami. Bioinformatics 27, 3315–3316.

Szumlanski, C., et al., 1996. Thiopurine methyltransferase pharmacogenetics: human gene cloning and characterization of a common polymorphism. DNA Cell Biol. 15, 17–30.

Tai, H., et al., 1997. 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. 94, 6444–6449.

Tsolis, A., et al., 2013. A consensus method for the prediction of aggregation-prone peptides in globular proteins. PLoS One 8 p. e54175.

Ujiie, S., et al., 2008. Functional characterization of 23 allelic variants of thiopurine S-methyltransferase gene (TPMT*2–*24). Pharmacogenet. Genomics 18, 887–893.

Van Durme, J., et al., 2009. Accurate prediction of DnaK-peptide binding via homology modelling and experimental data. PLoS Comput. Biol. 5, e1000475.

Wang, L., et al., 2003. Thiopurine S-methyltransferase pharmacogenetics: chaperone protein association and allozyme degradation. Pharmacogenetics 13, 555–564.

Wang, L., et al., 2005. Human thiopurine S-methyltransferase pharmacogenetics: Variant allozyme misfolding and aggresome formation. Proc. Natl. Acad. Sci. U. S. A. 102, 9394–9399.

Wang, L., et al., 2010. Very important pharmacogene summary: thiopurine S-methyltransferase. Pharmacogenet. Genomics 20, 401–405.

Waterhouse, A., et al., 2009. Jalview version 2—a multiple sequence alignment editor and analysis workbench. Bioinformatics 25, 1189–1191.

Weinshilboum, R., 2001. Thiopurine pharmacogenetics: clinical and molecular studies of thiopurine methyltransferase. Drug Metab. Dispos. 29, 601–605.

Wennerstrand, P., et al., 2012. Structural characteristics determine the cause of the low enzyme activity of two thiopurine S-methyltransferase allelic variants: a biophysical characterization of TPMT*2 and TPMT*5. Biochemistry 51, 5912–5920.

Whirl-Carrillo, M., et al., 2012. Pharmacogenomics knowledge for personalized medicine. Clin. Pharmacol. Ther. 92, 414–417.

Worth, C., Preissner, R., Blundell, T., 2011. SDM—a server for predicting effects of mutations on protein stability and malfunction. Nucleic Acids Res. 39, W215–22.

Wu, H., et al., 2007. Structural basis of allele variation of human Thiopurine S-Methyltransferase. Proteins 67, 198–208.

Yilmaz, A., 2015. Bioinformatic analysis of GJB2 gene missense mutations. Cell Biochem. Biophys. 71, 1623–1642.

Zaza, G., et al., 2010. Thiopurine pathway. Pharmacogenet. Genomics 20, 573–574.