Research Journal of Pharmaceutical, Biological and Chemical Sciences

Computational Analysis of Significant Missense Mutations in OAT gene that cause Gyrate Atrophy.

Jincy Anna Jomy, and Rao Sethumadhavan*

Bioinformatics Division, School of Bio Sciences and Technology, VIT University, Vellore, Tamil Nadu, India.

ABSTRACT

In this work the most detrimental missense mutations of OAT (Ornithine aminotransferase) that causes Gyrate Atrophy were identified computationally and the substrate binding efficiencies of those missense mutations were analyzed. The 22 missense mutations, I-Mutant 2.0, SIFT and PolyPhen programs were executed. In this we observed that 16 variants that were less stable, deleterious and damaging respectively. Subsequently, modeling of these 16 variants was performed to understand the change in their conformation with respect to the native OAT by computing their root mean square deviation (RMSD). Those 10 missense mutation namely, N54K, Y55H, N89K, C93F, R180T, A226V, P241L, T267I, A270P, R271K, H319Y, V332M, G353D, G375A, L402P and P417L were due to loss of stability in their mutant structures of OAT. This was confirmed by computing their total energies using GROMOS 96 force field and these mutations were cross validated with other well-known computational programs namely I-Mutant2.0, SIFT and PolyPhen-2.
Keywords: Missense mutation, Gyrate Atrophy, OAT, RMSD, Total energy, stabilizing residue

*Corresponding author

INTRODUCTION

Ornithine aminotransferase (OAT) (MIM: 258870) is a mitochondrial matrix enzyme that catalyzes the pyridoxal phosphate (PLP)-dependent transamination of ornithine to glutamic γ-semialdehyde [1]. OAT gene is located on the long (q) arm of chromosome 10 at position 26. More precisely, the OAT gene is located from base pair 126,085,871 to base pair 126,107,544 on chromosome 10 [1]. OAT deficiency (EC 2.6.1.13) is a rare congenital metabolic disorder characterized by gyrate atrophy of the choroid and retina causing a rare autosomal recessive disorder associated with two different clinical phenotypes [2]. Mutations that result in deficiency of OAT enzyme causes autosomal recessive eye disease called Gyrate Atrophy. The OAT gene encodes the mitochondrial enzyme ornithine aminotransferase, which is a key enzyme in the pathway that converts arginine and ornithine into the major excitatory and inhibitory neurotransmitters glutamate and GABA. Gyrate atrophy of the choroid and retina is an autosomal recessive chorioretinal dystrophy which leads to a slowly progressive loss of vision [3]. This disorder is characterized by night blindness, constriction of the field of vision, progressive peripheral retinal degeneration, an ultimately extinguished electroretinogram, and a marked elevation of plasma ornithine [4]. Other symptoms include vision loss, neonatal hyperammonemia (excess ammonia in the blood in the newborn period), neurological abnormalities, intellectual disability, peripheral nerve problems, and muscle weakness may occur. This condition is inherited in an autosomal recessive manner.

The level of OATase mRNA in the normal human retina is approximately equal to 1/100th the level of rhodopsin mRNA and 1/5th to 1/10th the level present in the retinoblastoma cells [5]. Ornithine aminotransferase (OAT) catalyzes a reaction in the pathway that interconvert ornithine and proline, and also serves to connect the urea and citric acid cycles. Loss of OAT function results in disease in humans, the enzyme is not a focus for rational inhibitor design. OAT is in the same subgroup as some other aminotransferases that also have key metabolic roles. This subgroup includes *γ*-aminobutyric acid aminotransferase (GABA-AT), an enzyme whose substrate is the brain's major inhibitory neurotransmitter, and glutamate-1-semialdehyde aminotransferase (GSA-AT), an essential enzyme in the tetrapyrrole synthesis pathway in plants [6]. Any information on the mechanism of inhibition with respect to OAT is relevant to rational drug design efforts on these other more attractive targets. L-canaline is virtually identical to ornithine. The crystal structure of OAT in complex with gabaculine provides the first structural evidence that the potency of the inhibitor is due to favorable aromatic–aromatic interactions with active-site residues. OATase is a pyridoxal phosphate-requiring enzyme that catalyzes the interconversion of ornithine, glutamate, and proline. Based on the crystal structure of human OAT, both substrate binding and reaction mechanism of the enzyme are well understood. OAT shows a large structural and mechanistic similarity to other enzymes from the subgroup III of aminotransferases, which transfer an amino group from a carbon atom that does not carry a carboxyl function [7]. Gyrate atrophy is obligate heterozygotes show approximately 50% of normal OAT activity, and consanguinity in parents of affected individuals is common. The mechanism by which the OAT deficiency leads to the choroidal and retinal atrophy and cataract formation remains unclear [8]. It is known that ornithine can be metabolized in three pathways (a) the conversion to glutamic acid γ-semialdehyde by OKT and then to proline or glutamate, (b) the conversion to citrulline by ornithine transcarbamylase in urea cycle and (c) the conversion to putrescine by ornithine decarboxylase (9). Ornithine delta aminotransferase controls the L-ornithine (Orn) level in tissues by catalysing the transfer of the delta amino group of Orn to 2-oxoglutarate. The products of this reaction are L-glutamate gama semialdehyde and L-glutamate. Among the compounds known to inhibit (or inactivate) OAT, only L-canaline and (SS)-5-(fluoromethyl)ornithine [(SS)-5FMOrn] are selective for OAT [9]. Identifying the disease-associated missense mutation had been a challenging task for genetic disorder research. Therefore, we attempted to investigate the mutants of OAT gene using a computational protocol that we devised for the analysis of BRCA1, CDKN2A, SMAD4 and ASPA [10–13]. The computational protocol was used to identify the detrimental missense mutations in ornithine aminotransferase and we proposed a model structure for the mutants.

MATERIALS AND METHODS

Datasets

From Swissprot/UniProt database, the protein sequence and variants (single amino acid polymorphisms/ missense mutations/point mutations) of ornithine aminotransferase were obtained. UniProtKB/Swiss-Prot is a manually curated knowledgebase providing information on protein sequences and

functional annotation. The 3D Cartesian coordinates of ornithine aminotransferase has been obtained from Protein Data Bank with PDB ID 1GBN [14]. A part of each Swissprot entry provides information on polymorphic variants, some of which polymorphic variants may be disease(s)- associated by causing defects in a given protein; most of them were nsSNPs (non-synonymous SNPs) in the gene sequence and SAPs (single amino acid polymorphisms) in the protein sequence[15-17].

Predicting stability changes caused by SAPs using support vector machine (I-Mutant 2.0)

I-Mutant2.0 (available at http://folding.uib.es/cgi-bin/i-mutant2.0.cgi) is a Support Vector Machine -based web server for the automatic prediction of protein stability changes upon single-site mutations. The tool was trained on a data set derived from ProTherm [18] that is presently the most comprehensive database of experimental data on protein mutations. Our predictor can evaluate the stability change upon single site mutation starting from the protein structure or from the protein sequence. When trained/tested with a cross validation procedure, I-Mutant2.0 correctly predicts whether the protein mutation stabilises or destabilises the protein in 80% of the cases when the three-dimensional structure is known and 77% of the cases when only the protein sequence is available. I-Mutant2.0 (available at http://folding.uib.es/cgi-bin/i-mutant2.0.cgi) is a support vector machine (SMV) based tool for the automatic prediction of protein stability changes caused by single point mutations. I-Mutant2.0 predictions were performed starting either from the protein structure or, more importantly, from the protein sequence [18]. This program was trained and tested on a dataset derived from ProTherm [19], which is the most comprehensive available database of thermodynamic experimental data of free energy changes of protein stability caused by mutations under different conditions. The output files show the predicted free energy change value or sign ($\Delta\Delta G$), which was calculated from the unfolding Gibbs free energy value of the mutated protein minus the unfolding Gibbs free energy value of the native protein (kcal mol^{-1}). Positive $\Delta\Delta G$ values meant that the mutated protein has higher stability and negative values indicate lower stability.

Analysis of functional consequences of point mutations by a sequence homology-based method (SIFT)

SIFT (available at http://sift.jcvi.org/www/SIFT_seq_submit2.html) program [20] is specifically used to detect deleterious single amino acid polymorphisms. SIFT is a sequence homology-based tool, which presumes that important amino acids will be conserved in a protein family; therefore, changes at well-conserved positions tend to be predicted as deleterious [21]. Here the queries are submitted in the form of protein sequences. SIFT takes a query sequence and uses multiple alignment information to predict tolerated and deleterious substitutions for every position of the query sequence. SIFT is a multistep procedure that, for given a protein sequence, (i) searches for similar sequences, (ii) chooses closely related sequences that may share similar function, (iii) obtains the multiple alignment of these chosen sequences, and (iv) calculates normalized probabilities for all possible substitutions at each position from the alignment. Substitutions at each position with normalized probabilities less than a chosen cutoff are predicted to be deleterious and those greater than or equal to the cutoff are predicted to be tolerated [20]. The cutoff value in SIFT program was tolerance index of ≥0.05. The higher the tolerance index, the less functional impact a particular amino acid substitution would be likely to have.

Simulation for functional change in a point mutant by structure homology-based method (PolyPhen-2)

We used the server PolyPhen-2 (at available http://genetics.bwh.harvard.edu/pph2/) for analyzing the damage caused by point mutations at the structural level is considered very important to understand the functional activity of the protein. We used the server PolyPhen-2 for this purpose [22]. Input options for the PolyPhen server are protein sequence, SWALL database ID or accession number, together with the sequence position of two amino acid variants. The query is submitted in the form of a protein sequence with a mutational position and two amino acid variants. Sequence-based characterization of the substitution site, profile analysis of homologous sequences, and mapping of the substitution site to known protein 3D structures are the parameters taken into account by PolyPhen-2 server to calculate the score. It calculates position-specific independent counts (PSIC) scores for each of the two variants and then computes the PSIC scores difference between them. The higher the PSIC score difference, the higher the functional impact a particular amino acid substitution would be likely to have.

Modeling Single Amino Acid Polymorphism (SAAP) location on protein structure to compute the RMSD

Structure analysis was performed to evaluate the structural deviation between native proteins and mutant proteins by means of root mean square deviation (RMSD). We used the web resource Protein Data Bank [17] and the single amino acid polymorphism database [23] (SAAPdb) to identify the 3D structure of OAT (PDB ID: 1GBN). We also confirmed the mutation position and the mutation residue in PDB ID. The mutation was performed *in silico* using the SWISSPDB viewer, and NOMAD-Ref server performed the energy minimization for 3D structures [24]. This server uses Gromacs as the default force field for energy minimization, based on the methods of steepest descent, conjugate gradient, and limited-memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) methods [25]. We used the conjugate gradient method to minimize the energy of the 3D structure of OAT. To optimize the 3D structure of OAT, we used the ifold server [26] for simulated annealing, which is based on discrete molecular dynamics and is one of the fastest strategies for simulating protein dynamics. This server efficiently samples the vast conformational space of biomolecules in both length and time scales. Divergence of the mutant structure from the native structure could be caused by substitutions, deletions and insertions [27] and the deviation between the two structures could alter the functional activity [28] with respect to binding efficiency of the inhibitors, which was evaluated by their RMSD values.

Computation of total energy and stabilizing residues

Total energy is one of the parameter that can indicate the stability between native and mutant modeled structures, and could be computed by the GROMOS96 force field that is embedded in the SWISSPDB viewer. Note that molecular mechanics or force field methods use classical type models to predict the energy of the molecule as a function of its conformation. This allows prediction of equilibrium geometries, transition states and relative energies between conformers or between different molecules. Molecular mechanics expresses the total energy as a sum of Taylor series expansions for the stretches for every pair of bonded atoms, and adds additional potential energy terms contributed by bending, torsional energy, van der Walls energy, and electrostatics [29]. Thus the total energy calculation could be considered as reliable parameter for understanding the stability of protein molecules with the aid of Force field (Gromos96 and Gromacs). Performing energy minimization and simulated annealing removes steric clashes and to obtains the best stable conformation [30]. Finally, the total energy was computed for native and mutant XRP2s by the GROMOS force field. Moreover, the total energy of the native structure was considered as a reference point for comparing the total energy of mutant structures for stability analysis. In addition, identifying the stabilizing residues for both the native and mutant structures represented a significant parameter for understanding their stability. Hence, we used the server SRide [31] to identify the stabilizing residues in the native and mutant protein models. Stabilizing residues were computed using parameters such as surrounding hydrophobicity, long-range order, stabilization center, and conservation score [31].

Calculating the total number of intra molecular interactions using PIC server

We used PIC server for computing intra-molecular interactions for both native and mutant structures respectively. PIC (Protein Interactions Calculator) server accepts atomic coordinate set of a protein structure in the standard Protein Data Bank (PDB) format. Interactions within a protein structure and interactions between proteins in an assembly are essential considerations in understanding molecular basis of stability and functions of proteins and their complexes. There are several weak and strong interactions that render stability to a protein structure or an assembly. It computes various interactions such as interaction between a polar residues, disulphide bridges, hydrogen bond between main chain atoms, hydrogen bond between main chain and side chain atoms, hydrogen bond between two side chain atoms, interaction between oppositely charged amino acids (ionic interactions), aromatic- aromatic interactions, aromatic- sulphur interactions and cation-π interactions. The PIC server [32] is available at: http://crick.mbu.iisc.ernet.in/~PIC.

RESULTS

Single Amino Acid Polymorphism Dataset from Swissprot

The 22 variants namely N54K, Y55H, N89K, Q90E, C93F, R154L, R180T, A226V, P241L, Y245C, R250P, T267I, A270P, R271K, H319Y, V332M, G353D, G375A, C394R, L402P, P417L and L437F investigated in this work were retrieved from Swissprot database (10-12).

Identification of functional variants by I-Mutant 2.0

Of the 22 variants, 17 variants were found to be less stable using the I-Mutant 2.0 server (Table 1) [18]. Among these 17 variants, one variant showed a ΔΔG value <-2.0. Seven variants showed a ΔΔG value <-1.0 and nine variants showed a ΔΔG value >-0.36 as depicted in Table 1. Of the 17 variants that showed a negative ΔΔG, three variants (P241L, T267I and P417L) changed from polar uncharged amino acid to non polar and three variants (A226V, V332M and G375A) has retain their amino acid properties. Two three variants (N54K and N89K) changed their polar uncharged amino acid to positively charged, two variant (R180T and R250P) changed from positively charged to polar uncharged and two variant (A270P and L402P) changed from non polar to polar uncharged. One variant (Y55H) changed from aromatic to positively charged, one variant (C93F) changed from polar uncharged to aromatic and one variant (R271K) remain positively charged. One variant (H319Y) changed positively charged to aromatic and one variant (G353D) changed non polar to negatively charged.Indeed, by considering only amino acid substitution based on physico-chemical properties, we could not be able to identify the detrimental effect. Rather, by considering the sequence conservation along with the above said properties could have more advantages and reliable to find out the detrimental effect of missense mutations [33].

Deleterious single point mutants identified by the SIFT program

The degree of conservation of a particular position in a protein was determined using sequence homology based tool SIFT [23]. The protein sequences of the 22 variants were submitted to SIFT to determine their tolerance indices. As the tolerance level increases, the functional influence of the amino acid substitution decreases and vice versa.

Among the 22 variants, 22 variants were found to be deleterious, having tolerance index scores of ≤0.05 (Table 1). Among these 22 variants, 19 variants showed a very high deleterious tolerance index score of 0.00. Two variant had a tolerance index score of 0.01 and one variant L437F had a tolerance index score of 0.05 (Table 1). Here, 17 deleterious variants identified by SIFT also were seen to be less stable by the I-Mutant 2.0 server.

Damaging single point mutations identified by the PolyPhen-2 server

Structural level alterations were determined by PolyPhen program. Protein sequence with mutational position and amino acid variants associated with the 14 single point mutants were submitted to the PolyPhen server [25]. A PSIC score difference of 0.5 and above was considered to be damaging. It could be seen from Table 1 that, all 22 variants were found to be damaging by PolyPhen-2. These variants also exhibited a PSIC score difference from 0.31 to 1. The whole 22 variants were found to be damaging by Polyphen were also deleterious by SIFT program. Similarly, out of the 22 variants, 17 variants are damaging I-Mutant 2.0.

Rational consideration of detrimental point mutations

We rationally considered the 17 most potential detrimental point mutations (N54K, Y55H, N89K, C93F, R180T, A226V, P241L, R250P, T267I, A270P, R271K, H319Y, V332M, G353D, G375A, L402P and P417L) for further course of investigations because they were commonly found to be less stable, deleterious, and damaging by the I-Mutant2.0, SIFT and Poly Phen-2 servers respectively [21,23,26]. We considered the statistical accuracy of these three programs, I-Mutant improves the quality of the prediction of the free energy change caused by single point protein mutations by adopting a hypothesis of thermodynamic reversibility of the existing experimental data. The accuracy of prediction for sequence and structure based values were 78% and 84% with correlation coefficient of 0.56 and 0.69, respectively [34]. SIFT correctly predicted 69% of the

substitutions associated with the disease that affect protein function. PolyPhen-2 evaluates rare alleles at loci potentially involved in complex phenotypes, densely mapped regions identified by genome-wide association studies, and analyses natural selection from sequence data, where even mildly deleterious alleles must be treated as damaging. PolyPhen-2 was reported to achieve a rate of true positive predictions of 92% [32-34]. To obtain precise and accurate measures of the detrimental effect of our variants, comprehensive parameters of all these three programs could be more significant than individual tool parameters. Hence, we further investigated these detrimental missense mutations by structural analysis. Figure 2 shows the list of functionally significant mutations with the commonly affected ones.

Table1: List of functionally significant mutants predicted to be less stable, deleterious and damaging by I-Mutant 2.0, SIFT and PolyPhen.

Variants	SIFT	PolyPhen	I-Mutant 2.0
N54K	**0**	**0.995**	**-0.79**
Y55H	**0**	**1**	**-1.92**
N89K	**0**	**0.993**	**-0.68**
Q90E	**0**	**1**	0.22
C93F	**0**	**0.997**	**-0.67**
R154L	**0**	**1**	0.19
R180T	**0**	**0.989**	**-1.89**
A226V	**0**	**0.828**	**-0.36**
P241L	**0**	**0.999**	**-1.07**
Y245C	**0**	**0.979**	0.25
R250P	**0.01**	0.444	**-0.64**
T267I	**0**	**1**	**-0.68**
A270P	**0.01**	**0.971**	**-1.63**
R271K	**0**	**1**	**-0.93**
H319Y	**0**	**1**	**-0.76**
V332M	**0**	**0.97**	**-1.89**
G353D	**0**	**1**	**-2.01**
G375A	**0**	**1**	**-1.11**
C394R	**0**	**0.501**	0.82
L402P	**0**	**1**	**-1.47**
P417L	**0**	**1**	**-0.6**
L437F	**0.05**	0.31	0.29

Notes: Letters in bold indicate mutants predicted to be less stable, deleterious and damaging by I-Mutant 2.0, SIFT and PolyPhen respectively.

Computing the RMSD by modeling of mutant structures

The available structure of OAT is PDB ID 1GBN. The mutational position and amino acid variants were mapped onto 1GBN native structure. Mutations at a specified position were performed *in silico* by SWISSPDB viewer independently to obtain a modeled structure. NOMAD-Ref server [19] and ifold server [26] performed the energy minimizations and stimulated annealing respectively, for both native structure and the 17 mutants modeled structures. To determine the deviation between the native structure and the mutants, we superimposed the native structures with all 17 mutant modeled structures and calculated the RMSD. The higher the RMSD value, the more deviation there is between the native and mutant structure, which in turn

changes the binding efficiency with the substrate because of deviation in the 3D space of the binding residues of OAT.

Table 2 shows the RMSD values for native structure with each mutant modeled structure. It shows that, two mutants exhibited a high RMSD >2.00 Å and one variant, 17 exhibited an RMSD >1.00 Å. Figure 3 shows the superimposed structures of native and mutants.

Table 2: RMSD, Total energy and Stabilizing residues for the native protein and mutants.

Variants	TE (Kj/mol^{-1})	SRIDE
Native	-19236	**ILE170**, **VAL171**, **ALA226**, **ILE261**, **GLY273**, **LEU289**, **GLY296**, **VAL304**, **LEU376**
Energy minimized Native	-21695	PHE44, ILE133, VAL134, ALA189, ILE224, ASP226, GLY236, ILE250, LEU252, GLY254, GLY259, VAL264, SER265, GLY288, ALA312, GLY338, ALA378
N54K	-21501	PHE81, **ILE170**, **VAL171**, **ALA226**, **ILE261**, ALA262, ASP263, **GLY273**, ILE287, **LEU289**, GLY291, **GLY296**, VAL301, SER302, GLY325, GLY375, ALA415
Y55H	-21391	PHE81, **ILE170**, **VAL171**, PHE172, **ALA226**, ILE224, ALA225, ASP226, GLY236, ILE287, **LEU289**, GLY291, **GLY296**, VAL301, SER302, GLY325, GLY375, LEU339, ALA415, PRO379
N89K	-21519	PHE81, GLN90, **ILE170**, **VAL171**, PHE135, **ALA226**, ILE224, ALA225, ASP226, ARG271, ILE287, **LEU289**, GLY291, **GLY296**, VAL301, SER302, GLY325, **LEU376**, ARG413, ALA415, LEU418
C93F	-21845	PHE81, **ILE170**, **VAL171**, PHE172, **ALA226**, ILE224, ALA225, ASP226, GLY236, ILE287, **LEU289**, GLY291, **GLY296**, VAL301, SER302, GLY325, GLY375, **LEU376**, ALA415, PRO416
R180T	-21469	PHE81, **ILE170**, **VAL171**, PHE172, **ALA226**, ILE224, ALA225, ASP226, GLY236, ILE287, **LEU289**, GLY291, **GLY296**, VAL301, SER302, GLY325, GLY375, **LEU376**, ALA415, PRO416
A226V	-21523	PHE81, **ILE170**, **VAL171**, VAL189, ILE224, ALA225, GLY236, ILE287, **LEU289**, LEU290, **GLY296**, VAL301, SER302, GLY375, **LEU376**, ALA415
P241L	-21914	PHE81, **ILE170**, **VAL171**, **ALA226**, PHE227, ILE224, ALA225, GLY236, ILE287, **LEU289**, GLY291, **GLY296**, VAL301, SER302, GLY325, GLY375, ALA415, PRO416
T267I	-21571	PHE81, **ILE170**, **VAL171**, **ALA226**, ILE224, ALA225, ASP226, GLY236, ILE287, **LEU289**, LEU290, GLY291, **GLY296**, VAL301, SER302, GLY325, GLY375, **LEU376**, ALA415
A270P	-21462	PHE81, **ILE170**, **VAL171**, PHE172, **ALA226**, ILE224, ALA225, ASP226, ILE287, **LEU289**, GLY291, **GLY296**, VAL301, SER302, GLY325, **LEU376**, ALA415
R271K	-21397	PHE81, **ILE170**, **VAL171**, **ALA226**, ILE224, ALA225, ASP226, ILE287, **LEU289**, **GLY296**, VAL301, SER302, ARG413, ALA415, PRO416
H319Y	-21991	PHE81, **ILE170**, **VAL171**, PHE172, **ALA226**, ILE224, ALA225, ASP226, GLY236, ILE287, **LEU289**, **GLY296**, VAL301, SER302, GLY325, GLY375, **LEU376**, ALA415
V332M	-21778	PHE81, **ILE170**, **VAL171**, **ALA226**, ILE224, ALA225, ASP226, GLY236, ILE287, **LEU289**, **GLY296**, VAL301, SER302, GLY375, ARG413, ALA415, PRO416
G353D	-21807	PHE81, **ILE170**, **VAL171**, **ALA226**, PHE227, ILE224, ALA225, ASP226, ILE287, **LEU289**, **GLY296**, VAL301, GLY325, ALA415, PRO416
G375A	-21991	PHE81, **ILE170**, **VAL171**, PHE172, **ALA226**, ILE224, ALA225, ASP226, ILE287, **LEU289**, GLY291, **GLY296**, VAL301, SER302, GLY325, **LEU376**, ALA415
L402P	-21929	PHE81, **ILE170**, **VAL171**, **ALA226**, PHE227, ILE224, ALA225, ASP226, ILE287, **LEU289**, GLY291, **GLY296**, VAL301, SER302, GLY325, **LEU376**, ALA415
P417L	-21914	PHE81, **ILE170**, **VAL171**, **ALA226**, PHE227, ILE224, ALA225, GLY236, ILE287, **LEU289**, GLY291, **GLY296**, VAL301, SER302, GLY325, GLY375, ALA415, PRO416

Notes: RMSD, root mean square deviation; the common stabilizing residues are shown in bold.

Application of GROMOS 96 and SRIDE for native structure and mutant modeled structures

The total energy was calculated for both native and mutant structures. Table 2 shows that total energy of native structure was –19236 kcal mol^{-1}. Whereas the 11 mutant structures all had slightly higher total energies compared with the native structure. Note that the higher the total energy, the lesser the stability and vice versa. We then used the SRide server to identify the stabilizing residues of both the native structure and the mutant modeled structures (Table 2). The native structure has 17 stabilizing residues whereas on the other hand, the mutant structures have between 15 and 21 stabilizing residues. This clearly indicates that 17 mutant structures were less stable than the native structure.

Computing the intra-molecular interactions in OAT

We further validated the stability of protein structure by using the PIC server [30] to identify the number of intra-molecular interactions for both native and mutant structures (Table 3). Interactions within a

protein structure and the interactions between proteins in an assembly were essential considerations in understanding molecular basis of stability and functions of proteins and their complexes. There were several weak and strong intra-molecular interactions that render stability to a protein structure. Therefore these intra-molecular interactions were computed by PIC server in order to further substantiate the stability of protein structure. Based on this analysis, we found that a total number of 1434 intra-molecular interactions were obtained in the native structure of OAT. On the other hand, 17 mutant structures of RP2 established the intra-molecular interactions between the range of 1404 to 1479 as shown in Table 3.

Table 3: It shows the no: of Intra-molecular interactions of the native protein and mutants

Variants	Total	HI	MM	MS	SS	II	AA	AS	CI
Native	1255	381	517	174	119	34	12	4	14
EM_Native	1434	384	570	251	160	33	13	6	17
N54K	1442	379	569	246	171	35	16	7	19
Y55H	1454	382	572	246	180	38	12	8	16
N89K	1411	381	563	253	142	35	13	8	16
C93F	1449	383	566	252	175	35	13	8	17
R180T	1404	380	562	235	159	32	12	8	16
A226V	1416	376	552	244	178	33	12	6	15
P241L	1442	387	568	247	167	37	13	7	16
T267I	1413	382	560	237	164	33	13	7	17
A270P	1464	390	565	257	176	36	14	7	19
R271K	1427	381	562	257	156	32	14	8	17
H319Y	1465	392	564	259	175	34	13	9	19
V332M	1429	379	563	256	159	34	15	6	17
G353D	1446	380	574	264	157	34	13	7	17
G375A	1464	399	565	260	169	34	13	8	16
L402P	1479	386	584	268	169	37	14	5	16
P417L	1442	387	568	247	167	37	13	7	16

Notes: Total no of intramolecular interactions
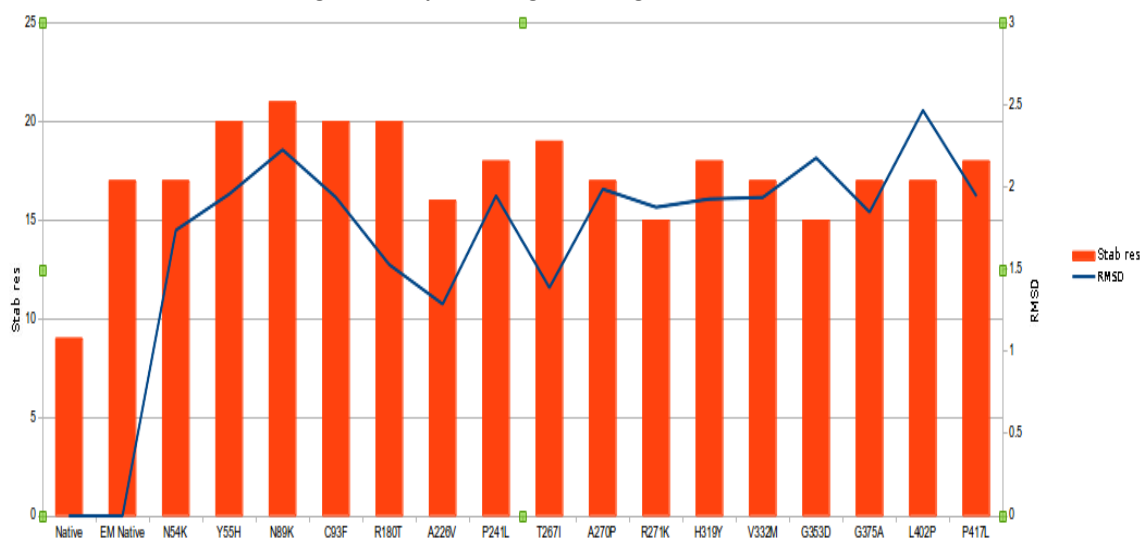
Figure 1: Graph showing stabilizing residue with RMSD.

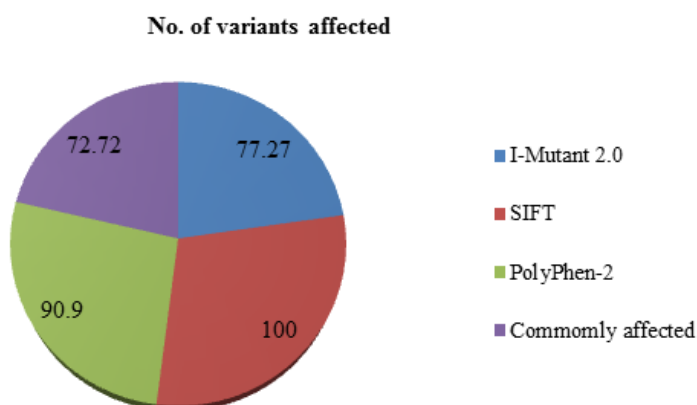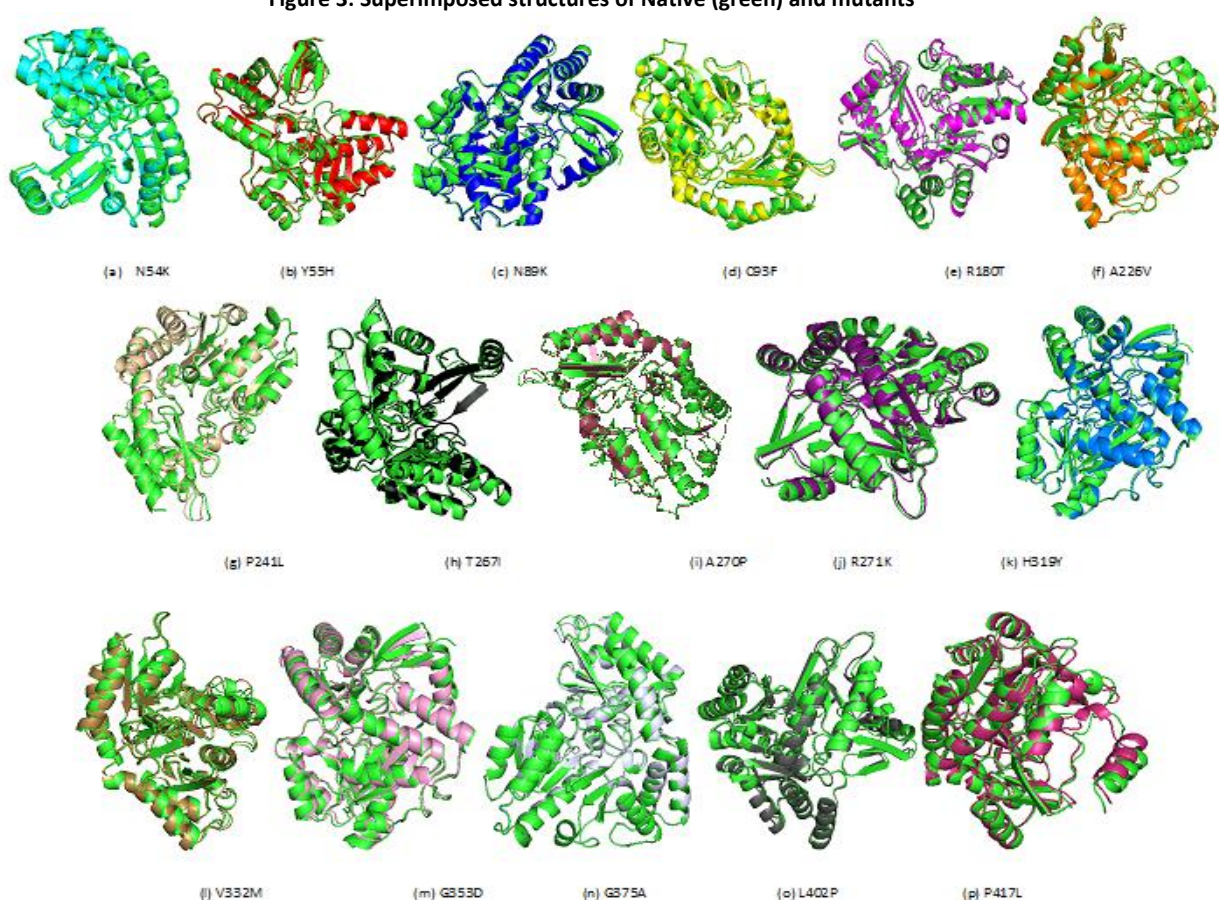**Figure 2: List of functionally significant mutations.**

No. of variants affected

- I-Mutant 2.0
- SIFT
- PolyPhen-2
- Commomly affected

Figure 3: Superimposed structures of Native (green) and mutants

(a) N54K (b) Y55H (c) N89K (d) C93F (e) R180T (f) A226V

(g) P241L (h) T267I (i) A270P (j) R271K (k) H319Y

(l) V332M (m) G353D (n) G375A (o) L402P (p) P417L

(a) Superimposed structure of native 1GBN (green) with mutant N54K (cyan) structure showing RMSD of 1.74Å. (b) Superimposed structure of native 1GBN (green) with mutant Y55H (red) structure showing RMSD of 1.96Å. (c) Superimposed structure of native 1GBN (green) with mutant N89K (blue) structure showing RMSD of 2.23Å. (d) Superimposed structure of native 1GBN (green) with mutant C93F (yellow) structure showing RMSD of 1.94Å. (e) Superimposed structure of native 1GBN (green) with mutant R180T (magenta) structure showing RMSD of 1.53Å. (f) Superimposed structure of native 1GBN (green) with mutant A226V (orange) structure showing RMSD of 1.29Å. (g) Superimposed structure of native 1GBN (green) with mutant P241L (wheat) structure showing RMSD of 1.95Å. (h) Superimposed structure of native 1GBN (green) with mutant T267I (black) structure showing RMSD of 1.39Å. (i) Superimposed structure of native 1GBN (green) with mutant A270P (raspberry) structure showing RMSD of 1.99Å. (j) Superimposed structure of native 1GBN (green) with mutant R271K (purple) structure showing RMSD of 1.88Å. (k) Superimposed structure of native 1GBN (green) with mutant H319Y (skyblue) structure showing RMSD of 1.93Å. (l) Superimposed structure of native 1GBN (green) with mutant V332M (sand) structure showing RMSD of 1.94Å. (m) Superimposed structure of native 1GBN (green) with mutant G353D (lightpink) structure showing RMSD of 2.18Å. (n) Superimposed structure of native 1GBN (green) with mutant G375A (bluewhite) structure showing RMSD of 1.85Å. (o) Superimposed structure of native 1GBN (green) with mutant L402P (gray) structure showing RMSD of 2.47Å. (p) Superimposed structure of native 1GBN (green) with mutant P417L (hotpink) structure showing RMSD of 1.95Å.

CONCLUSION

Of the 22 variants that were retrieved from Swissprot, 22 variants were found less stable by I-PolyPhen and SIFT and 17 variants were considered damaging by Mutant 3.0. Seventeen variants were selected as potentially detrimental point mutations because they were commonly found to be less stable, deleterious and damaging by the I-Mutant 2.0, SIFT and Poly-Phen servers, respectively. The structures of these 17 variants were modeled and the RMSD between the mutants and native structures ranged from 1.29Å to 2.47Å. Finally, we concluded that the lower binding affinity of 17 mutants (N54K, Y55H, N89K, C93F, R180T, A226V, P241L, R250P, T267I, A270P, R271K, H319Y, V332M, G353D, G375A, L402P and P417L) of their Free energy and RMSD scores identified them as deleterious mutations. Thus the results indicate that our approach successfully allowed us to (1) consider computationally a suitable protocol for missense mutation (point mutation/single amino acid polymorphism) analysis before wet lab experimentation and (2) provided an optimal path for further clinical and experimental studies to characterize mutants in depth.

ACKNOWLEDGMENT

REFERENCES

[1]   Mara Doimo, Maria Andrea Desbats, Maria Cristina Baldoin, Elisabetta Lenzini, Giuseppe Basso, Elaine Murphy, Graziano C, Seri M, Burlina A, Sartori G, Trevisson E, Salviati L. Hum Mutat 2013; 34: 229–236.
[2]   Ohkubo Y, Ueta A, Ito T, Sumi S, Yamada M and Ozawa K, Togari H. Tohoku J Exp Med 2005; 205: 335-42.
[3]   Santinelli R, Costagliola C, Tolone C, D'Aloia A, D'Avanzo A and Prisco F, Perrone L, del Giudice EM. J Inherit Metab Dis. 2004; 27:187-96.
[4]   James J. O'Donnell, Kaarina Vannas-Sulonen, Thomas Shows B, David R. Cox. Am J Hum Genet 1988; 43: 922-928.
[5]   Inana G, Totsuka S, Redmond M, Dougherty T, Nagle J, Shiono T, Ohura T, Kominami E, Katunuma N. Proc Natl Acad Sci U S A 1986; 83: 1203-7.
[6]   Sapan A Shah, Betty Shen W, Axel Brünger T. Structure 1997; 5: 1067-1075.
[7]   Jana Stranska, David Kopecny, Martina Tylichova, Jacques Snegaroff, Marek Sebela. Plant Signal Behav 2008; 3: 929–935.
[8]   Dovid J Barrett, J Bronwyn Bateman, Robert Sparkes S, Mohandas T, Ivana Klisak F, George lnono. Invest Ophth Vis Sci 1987; 28: 1037-42.
[9]   Seiler N. Curr Drug Targets. 2000; 1: 119-53.
[10]   Rajasekaran R, Sudandiradoss C, Doss CG, Sethumadhavan R. Genomics 2007; 90: 447–452.
[11]   Rajasekaran R, Priya Doss CG, Sudandiradoss C, Ramanathan K, Sethumadhavan R. Biochimie 2008; 90: 1523–1529.
[12]   Rajasekaran R, Sethumadhavan R. J Bionanosci 2009; 80-87.
[13]   Sreevishnupriya K, Chandrasekaran P, Senthilkumar A, Sethumadhavan R, Shanthi V, Daisy P, Nisha J, Ramanathan K, Rajasekaran R. Sci China Life Sci 2012; 55: 1190-1119.
[14]   Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. Nucleic Acids Res 2000; 28: 235–242.
[15]   Yip YL, Scheib H, Diemand AV, Gattiker A, Famiglietti LM, Gasteiger E, Bairoch A. Hum Mutat 2004; 23: 464–470.
[16]   Yip YL, Famiglietti M, Gos A, Duek PD, David FP, Gateau A, Bairoch A. Hum Mutat 2008; 29: 361–366.
[17]   Brigitte Boeckmann, Amos Bairoch, Rolf Apweiler, Marie-Claude Blatter, Anne Estreicher, Elisabeth Gasteiger, Martin MJ, Michoud K, O'Donovan C, Phan I, Pilbout S, Schneider M. Nucleic Acids Res 2003; 31: 365–370.
[18]   Capriotti E, Fariselli P, Casadio R. Nucleic Acids Res 2005; 33: 306–310.
[19]   Bava KA, Gromiha MM, Uedaira H, Kitajima K, Sarai A. Nucleic Acids Res 2004; 32: 120–121.
[20]   Ng PC, Henikoff S. Nucleic Acids Res 2003; 31: 3812–3814.
[21]   Ng PC, Henikoff S. Genome Res 2001; 11: 863–874.
[22]   Ramensky V, Bork P, Sunyaev S. Nucleic Acids Res 2002; 30: 3894–3900.

[23]   Cavallo A, Martin AC, Bioinformatics 2005; 21: 1443–1450.
[24]   Lindahl E, Azuara C, Koehl P, Delarue M. Nucleic Acids Res 2006; 34: 52–56.
[25]   Delarue M, Dumas P. P Natl Acad Sci USA 2004; 101: 6957–6962.
[26]   Sharma S, Ding F, Nie H, Watson D, Unnithan A, Lopp J, Pozefsky D, Dokholyan NV. Bioinformatic 2006; 22: 2693–2694.
[27]   Han JH, Kerrison N, Chothia C, Teichmann SA. Structure 2006; 14: 935–945.
[28]   Varfolomeev SD, Uporov IV, Fedorov EV. Biochem 2002; 67: 1099–1108.
[29]   Leach AR. Molecular Modeling: Principles and Applications. Pearson Education EMA, Sussex 2001.
[30]   Chou KC, Carlacci L. Protein Eng 1991; 4: 661–667.
[31]   Magyar C, Gromiha MM, Pujadas G, Tusnády GE, Simon I. Nucleic Acids Res 2005; 33: 303–305.
[32]   Tina KG, Bhadra R, Srinivasan N. Nucleic Acids Res 2007; 35: 473-476.
[33]   Connolly ML. Science 1983; 221: 709–713.
[34]   Duhovny D, Nussinov R, Wolfson Editors HJ Lecture Notes in Computer Science, Springer 2002; 2452: 185–200.