ISSN: 0975-8585

Research Journal of Pharmaceutical, Biological and Chemical Sciences

Computational Identification and Structural Analysis of Deleterious Functional SNPs in ARL6 Gene Causing Bardet-Biedl Syndrome.

Jincy Anna Jomy, and Rao Sethumadhavan*.

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

ABSTRACT

Here we have evaluated the Single Nucleotide Polymorphisms (SNPs) that can alter the expression and the function in ARL6 gene through computational methods. To explore possible relationships between genetic mutations and phenotypic variation, different computational methods like Sorting Intolerant from Tolerant (SIFT, an evolutionary-based approach), Polymorphism Phenotyping (PolyPhen, a structure-based approach) and I-Mutant 3.0 (support vector machine based tool) are discussed. There were 5 missense mutations; in this we observed 4 variants that were deleterious and damaging respectively. We got all 5 non-synonymous SNPs (nsSNPs) (100%) to be deleterious by SIFT and I- Mutant 3.0 and 4 nsSNPs (80%) as from Polyphen-2. Then cation-π interactions in protein structures are identified and analyzed the roles played by Arg, Lys interactions with π (Phe, Tyr or Trp) residues and their role in the structural stability. Subsequently, modeling of these 5 variants was performed to understand the change in their conformation with respect to the native ARL6 by computing their root mean square deviation (RMSD). Those 4 missense mutations were due to loss of stability in their mutant structures of ARL6. This was confirmed by computing their total energies using GROMOS 96 force field and these mutations were cross validated with computational programs.

Keywords: Missense mutation, ARL6, RMSD, Total energy, stabilizing residue, π- interactions

Corresponding author

ISSN: 0975-8585

INTRODUCTION

Bardet- Biedl syndrome (BBS) is an autosomal recessive condition characterised by rod-cone dystrophy, postaxial polydactyly, central obesity, mental retardation, hypogonadism, and renal dysfunction [1]. BBS is a disorder that affects many parts of the body. The signs and symptoms of this condition vary among affected individuals, even among members of the same family. Vision loss is one of the major features of BBS. Loss of vision occurs as the light-sensing tissue at the back of the eye (the retina) gradually deteriorates. Almost 12 genes have been identified to be responsible for BBS so far [2]. Although the BBS proteins show little structural similarity, a set of BBS proteins are located in the neighborhood of cilia and are involved in the ciliogenesis and/or cilia function[3]. ARL6 (ADP-ribosylation factor-like 6) is the protein-coding gene which belongs to the Arf/Arl-family GTPases has shown to be a cause for BBS using comparative genomic analysis [4]. Bardet-biedl syndrome 1 and arl6-related retinitis pigmentosa are other diseases that associated with ARL6. It involves in membrane protein trafficking at the base of the ciliary organelle and mediates recruitment onto plasma membrane of the BBSome complex which would constitute a coat complex required for sorting of specific membrane proteins to the primary cilia [5]. There is significant phenotypic and molecular overlap between Bardet-Biedl syndrome and other cilopathies [6].

Mutations in several genes that cause BBS can also lead to other distinct ciliopathy syndromes. The main clinical features are cone-rod dystrophy, with childhood-onset vision loss preceded by night blindness; postaxial polydactyly; truncal obesity that manifests during infancy and remains problematic throughout adulthood; intellectual disability; male hypogenitalism and complex female genitourinary malformations; and renal dysfunction, which is a major cause of morbidity and mortality [1] BBS expression varies both within and between families and diagnosis is often difficult [6].

Hence, in this present study, we performed computational analysis of nsSNPs in ARL6 gene and identified the possible deleterious mutations by SIFT, PolyPhen and I-Mutant 3.0. In addition, structural analysis was performed to identify the π interactions and intra molecular interactions of native and mutants to evaluate the protein stability.

MATERIALS AND METHODS

Datasets

The protein sequence and variants (single amino acid polymorphisms/ missense mutations/point mutations) of ARL6 were obtained from the Swissprot database. The 3D Cartesian coordinates of ADP-ribosylation factor-like protein 6 (ARL6) has been obtained from Protein Data Bank with PDB ID 2H57 which has a resolution of 2.00Å [7-10].

Analysis of functional consequences of coding nsSNPs by sequence homology based method (SIFT)

SIFT (http://sift.jcvi.org/www/SIFT_seq_submit2.html) [11] is used to detect the deleterious coding non-synonymous SNPs. SIFT is a sequence homology-based tool that presumes that important amino acids will be conserved in the protein family. Hence, changes at well-conserved positions tend to be predicted as deleterious [12]. The queries have been submitted as protein sequence. The principle of this program is that it 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, given a protein sequence, (a) searches for similar sequences, (b) chooses closely related sequences that may share similar functions, (c) obtains the multiple alignment of the chosen sequences, and (d) 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 [11]. SIFT score ranges from 0 to 1. The amino acid substitution is predicted damaging is the score is <= 0.05, and tolerated if the score is > 0.05.

Simulation for functional change in coding nsSNPs by structure homology based Method (PolyPhen)

PolyPhen-2 (http://genetics.bwh.harvard.edu/pph2/) [13] is used to analyze the damage caused by point mutations at the structural level is considered very important to understand the functional activity of the

protein. The query is submitted in the form of a protein sequence with a mutational position and two amino acid variants. Structural level alterations were determined by PolyPhen program. 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.

Predicting stability changes caused by SAPs using support vector machine (I-Mutant 3.0)

I-Mutant 3.0 (http://folding.uib.es/cgi-bin/i-mutant3.0.cgi) is a suite of Support Vector Machine (SVM) based predictors integrated in a unique web server which offers the opportunity to predict automatically protein stability changes upon single-site mutations starting from protein sequence alone or protein structure when available. This program gives the possibility to predict human deleterious Single Nucleotide Polymorphism starting from the protein sequence [14]. We used the sequence-based version of I-Mutant 3.0 that classifies the prediction in three classes: neutral mutation ($-0.5 \leq DDG \leq 0.5$ kcal/mol), large decrease (≤ -0.5 kcal/mol) and large increase (> 0.5 kcal/mol). This program was trained and tested on a dataset derived from ProTherm [15], 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.

Computation of Cation–π Interactions Energy

Cation-π interactions in protein structures are identified and evaluated by using an energy-based criterion for selecting significant sidechain pairs [16]. These cation–π interactions are obtained using CaPTURE (http://capture.caltech.edu/) program. Cation-π interactions are found to be common among structures in the Protein Data Bank. . The energies are computed for all the pairs of cationic-aromatic amino acid residues (Arg/Lys with Phe, Tyr and Trp). The total Cation–π interaction energy (Ecat–π) has been divided into electrostatic (Ees) and van der Waals energy (Evw) and was computed using the program CaPTURE, which had implemented a subset of OPLS force field21 to calculate the energies. The Ecat-π is the sum of these two energies, i.e., electrostatic and the van der Waal's energy [17].

$$E_{cat-\pi} = E_{es} + E_{vdW}$$

Conservation score calculation.

The conservation score of residues were calculated using Consurf server (http://consurf.tau.ac.il/). Conservation score is a useful parameter for the identification of conserved residues in a protein sequence [18].

Secondary Structure and Solvent Accessibility

We obtained the information about secondary structures DSSP (http://www.cmbi.ru.nl/dssp.html) [19]. The secondary structures were classified into α-helix, β-strand, and random coil [20]. We obtained the solvent accessible surface area (ASA) information using Net ASA view (http://www.netasa.org/). Solvent accessibility was divided into three classes: buried, partially exposed, and exposed indicating respectively the least, moderate, and high accessibility of the amino acid residues to the solvent [21]. In order to obtain the preference and pattern of each cation-π interaction-forming residue in glycoproteins, we conducted a systematic and careful analysis based on their location in different secondary structures and their solvent accessibility [22].

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 [10] and the single amino acid polymorphism database [23] (SAAPdb) to identify the 3D structure of ARL6 (PDB ID: 2H57). We also confirmed the mutation position and the mutation residue in PDB ID 2H57. The mutation was performed *in silico* using the SWISSPDB viewer, and NOMAD-Ref server performed the energy minimization for 3D structures. 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 [24]. We used the conjugate gradient method to minimize the energy of the 3D structure of ARL6. To optimize the 3D structure of ARL6, we used the ifold server 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 [25]. Divergence of the mutant structure from the native structure could be caused by substitutions, deletions and insertions [26] and the deviation between the two structures could alter the functional activity [27] with respect to binding efficiency of the inhibitors, which was evaluated by their RMSD values.

Computation of total energy

Total energy could be computed by the GROMOS96 force field that is embedded in the SWISSPDB viewer. 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 [28]. 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 [29]. Finally, the total energy was computed for native and mutant ARL6s by the GROMOS force field.

Computation of stabilizing residues

To check the stability for the native and mutant modeled structures, identification of the stabilizing residues will be useful. Total energy is one of the parameter that can indicate the stability between native and mutant modeled structures. 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 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 [30].

Calculating the total number of intra molecular interactions using PIC server

Protein Interactions Calculator (http://crick.mbu.iisc.ernet.in/~PIC) server is used for computing intra-molecular interactions for both native and mutant structures. PIC 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 [31].

ISSN: 0975-8585

RESULTS AND DISCUSSION

Single Amino Acid Polymorphism Dataset from Swissprot

The 5 variants namely T31M, T31R, A89V, G169A and L170W investigated in this work were retrieved from Swissprot database [7-9].

Damaging single point mutations identified by the PolyPhen-2 server

Protein sequence with mutational position and amino acid variants associated with the 5 single point mutants were submitted to the PolyPhen server and 4 variants were found to be damaging. A PSIC score difference of 0.5 and above was considered to be damaging. These variants also exhibited a PSIC score difference from 0.778 to 1(Table 1) [13].

Table 1: List of functionally significant mutants predicted to be by I-Mutant 3.0, SIFT and PolyPhen-2

Variants	SIFT	PolyPhen 2	I-Mutant 3.0
T31M	**0**	**1**	**-0.39**
T31R	**0**	**1**	**-0.62**
A89V	**0.03**	0.778	**-0.1**
G169A	**0**	**1**	**-0.47**
L170W	**0**	**1**	**-1.07**

Notes: Letters in bold indicate mutants predicted to be less stable, deleterious and damaging by I-Mutant 3.0, SIFT and PolyPhen-2 respectively.

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. The protein sequences of the 5 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. All the 5 variants were found to be deleterious with tolerance index scores of ≤0.05 (Table 1). Among these 5 variants, 4 variants showed a very high deleterious tolerance index score of 0.00. One variant A89V had a tolerance index score of 0.03 (Table 1) [11]. The variants were found to be damaging by SIFT were also deleterious by PolyPhen program and Mutant 3.0.

Identification of functional variants by I-Mutant 3.0

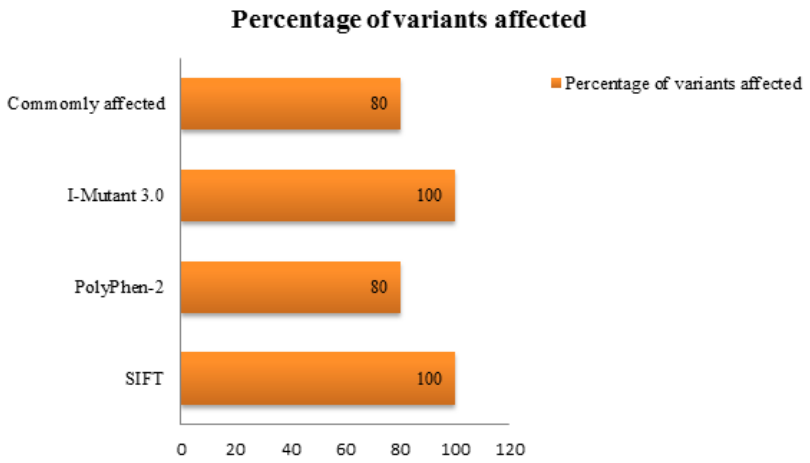
Of the 5 variants, all were found to be less stable using the I-Mutant 3.0 server (Table 1) [14]. Among these 5 variants, one variant showed a $\Delta\Delta G$ value <-1.07, one variants showed a $\Delta\Delta G$ value <-0.1, the other variant showed a $\Delta\Delta G$ value >-0.62 as depicted in Table 1. Of the 5 variants that showed a negative $\Delta\Delta G$, one variant (T31M) changed polar uncharged amino acid to non polar amino acid and one variant (T31R) changed polar uncharged to positively charged, one variant (L170W) changed non polar amino acid to aromatic and two variants (A89V and G169A) remain as non polar amino acid. 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 [32].

Rational consideration of detrimental point mutations

We have considered the 4 most potential detrimental point mutations (T31M, T31R, G169A and L170W) for further course of investigations because they were commonly found to be less stable, deleterious and damaging by the I-Mutant3.0, SIFT and Poly Phen-2 servers respectively [14,11,13]. 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 [30]. 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% [28-30]. 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 1 shows the list of functionally significant mutations with the commonly affected ones.

Figure 1: List of functionally significant mutations



Computing the RMSD by modeling of mutant structures

The available structure of ARL6 is PDB ID 2H57. The mutational position and amino acid variants were mapped onto 2H57 native structure. Mutations at a specified position were performed *in silico* by SWISSPDB viewer independently to obtain a modeled structure. NOMAD-Ref server [15] and ifold server [25] performed the energy minimizations and stimulated annealing respectively, for both native structure and the 4 mutants modeled structures. To determine the deviation between the native structure and the mutants, we superimposed the native structures with all 4 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 ARL6. Table 2 shows the RMSD values for native structure with each mutant modeled structure. Table 2 shows that, one of the mutant (L170W) exhibits a high RMSD >1.00 Å. Figure 3 shows the pymol view of superimposed structures of native and mutants.

Table 2: RMSD, total energy and stabilizing residues for the native protein and mutants.

Variants	RMSD	Total energy (Kj/mol)	No. of SR	Sride
Native		-9896.673	16	VAL20, LEU21, CYS22, LEU23, THR66, VAL67, PHE68, ALA89, ILE91, PHE92, PRO124, LEU126, PHE127, PHE128, ALA129, LYS155
T31M	0.61Å	-9829.93	15	**VAL20, LEU21, CYS22, LEU23**, LYS57, **THR66, VAL67, PHE68,** MET70, **ALA89, ILE91, PHE92**, ILE123, **LEU126, ALA129**
T31R	0.67Å	**-9992.546**	13	**VAL20, LEU21, CYS22, LEU23, THR66, VAL67, PHE68,** MET70, **ALA89, ILE91, PHE92, LEU126, ALA129**
G169A	0.55Å	-9729.935	14	**VAL20, LEU21, CYS22, LEU23, THR66, VAL67, PHE68,** MET70, **ALA89, ILE91, PHE92**, ILE123, **LEU126, ALA129**
L170W	1.44Å	**-9974.515**	**17**	**VAL20, LEU21, CYS22, LEU23,** SER54, LYS57, PHE58, **THR66, VAL67,** ASP69, MET70, **ALA89, ILE91, PHE92,** VAL93, **LEU126, ALA129**

Notes: RMSD- Root Mean Square Deviation; SR- Stabilizing residues; the common stabilizing residues are shown in bold

ISSN: 0975-8585

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 -9896.673 kcal mol^{-1}, whereas the 4 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). But the native and mutants shows no stable structures.

Computing the intra-molecular interactions in ARL6

We further validated the stability of protein structure by using the PIC server [26] 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 1392 intra-molecular interactions were obtained in the native structure of 2H57. On the other hand, 4 mutant structures is established the intra-molecular interactions between the range of 594 to 626 as shown in Table 3.

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

Variants	Total	HI	DB	MM	MS	SS	II	AA	AS	CI
Native_2H57	1392	429	0	576	576	102	44	18	8	2
T31M	594	147	0	232	125	60	23	5	1	1
T31R	597	149	0	227	130	58	26	5	1	1
G169A	601	148	0	228	127	67	24	5	1	1
L170W	626	145	0	234	136	79	24	7	0	1

Notes: Total no of intramolecular interactions. HI- Hydrogen Interractions, MM- Main chain-Main chain interaction, MS- Main chain Side chain interaction, SS- Side chain side chain interactions, II- Ionic-Ionic interaction, AA- Aromatic-Aromatic interactions, AS- Aromatic-Sulphur interactions, CI- Cation-π interactions

Energetically Significant Cation–π Interactions.

The Cation– π interaction energy of both native and mutant prion proteins was analysed. The two pairs of cation-π interactions (Arginine-Tyrosine and Arginine- Tryptophan) in native are 3.62 and 3.96 respectively(Table 4). On the other hand mutants show 3.62 and 3.98, which has least energy which shows slightly lowest cation-π interactions. So this indicates four mutants established slightly strong cation-π interaction than Native [16]. The results are shown in Figure 2.

Table 4: Average cation-π interaction energy

Variants	**R-F (-Kcal/mol)**	**R-Y (-Kcal/mol)**	**R-W (-Kcal/mol)**	**K-F (-Kcal/mol)**	**K-Y (-Kcal/mol)**	**K-W (-Kcal/mol)**
Native_ 2H57	-	R75-Y76(3-.62)	R77-W80(3.96)	-	-	-
T31M	-	R75-Y76 (-3.62)	R77-W80 (-3.98)	-	-	-
T31R	-	R75-Y76 (-3.62)	R77-W80 (-3.98)	-	-	-
G169A	-	R75-Y76 (-3.62)	R77-W80 (-3.98)	-	-	-
L170W	-	R75-Y76 (-3.62)	R77-W80 (-3.98)	-	-	-

Notes: R- Arginine, F- Phenylalanine, Y- Tyrosine, W- Tryptophan, K - Lysine

ISSN: 0975-8585

Figure 2 : Composition of cation-π residues

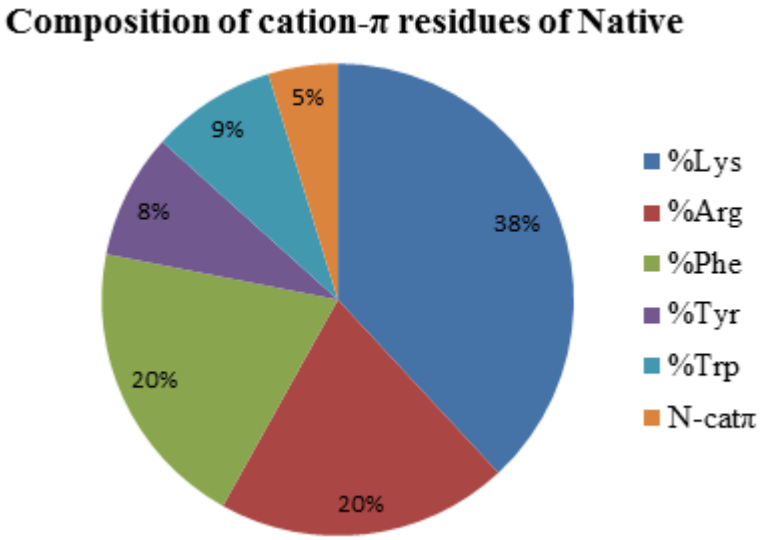
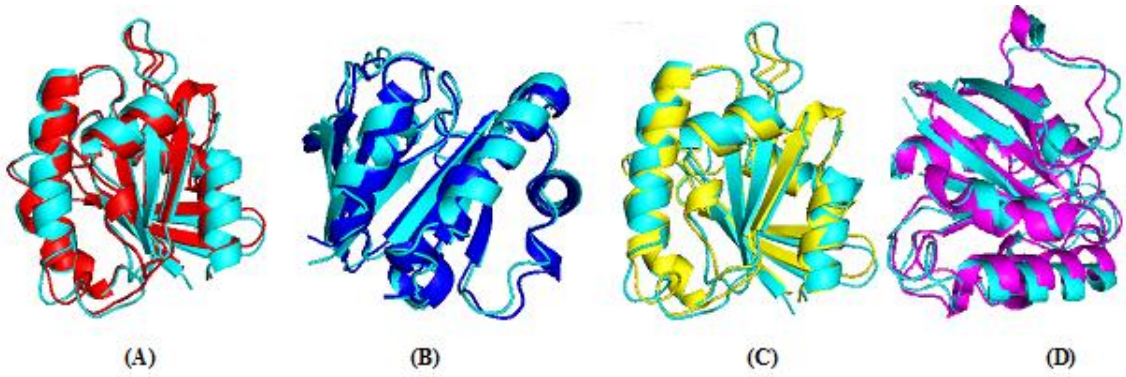


Figure 3. Superimposed structure of the native protein (cyan) with mutants



(A) Superimposed structure of native ARL6 (cyan) with mutant T31M (red) structure showing RMSD of 0.61Å. (B) Superimposed structure of native ARL6 (cyan) with mutant T31R (blue) structure showing RMSD of 0.67Å. (C) Superimposed structure of native ARL6 (cyan) with mutant G169A (yellow) structure showing RMSD of 0.55Å. (D) Superimposed structure of native ARL6 (cyan) with mutant L170W (magenta) structure showing RMSD of 1.44Å.

Secondary structure Preferences

The occurrence of weak interactions has been observed at the terminus of the secondary structural units, in particular α-helix and β-sheets [33, 34]. These interactions play a definitive role in stabilizing the proteins. Here we have calculated the occurrence of cation-π forming residues in secondary structure. We found that the cation-π forms Turns (T) and 3_{10} helix (G) which is shown in Table 5.

Table 5: Secondary structure Preferences

PDB ID	Cat-residue	ASA	2° str	π-residue	2° str	ASA	D_{seq}
2H57	R75	90	T	Y76	T	25	1
	R77	28	G	W80	G	2	3

Notes: T- Turns, G- 310 helix

CONCLUSION

In our analysis, we identified the most deleterious mutation in ARL6 based on various in silico tools. Of 5 variants that were retrieved from Swissprot, all 5 were found detrimental by SIFT and I-Mutant 3.0 respectively and 4 variants were found less stable by PolyPhen. As four variants were commonly found to be less stable, deleterious and damaging by SIFT, PolyPhen and I-Mutant 3.0, they were selected as potentially detrimental point mutations. The structures of these 4 variants were modeled and the RMSD between the mutants and native structures were calculated. The cation-π interacting residues are found to stabilize both the regular and non-regular secondary structural elements. Arginine-Tyrosine and Arginine- Tryptophan pairs have strongest cation-π interactions. Finally we concluded that the lower binding affinity of 4 mutants (T31M, T31R, A89V, G169A and L170W) 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 BBS mutants in depth.

ACKNOWLEDGMENT

The authors thank the management of Vellore Institute of Technology University for providing the facilities to carry out this work.

REFERENCES

[1] Elizabeth Forsythe, Philip L Beales. Eur J Hum Genet 2013; 21: 8–13.
[2] Slavotinek AM, Stone EM, Mykytyn K, Heckenlively JR, Green JS, Heon E, Musarella MA, Parfrey PS, Sheffield VC, Biesecker LG. Nat Genet 2000; 26:15-6.
[3] Alok Shah S, Sara Farmen L, Thomas Moninger O, Thomas Businga R, Michael Andrews P, Kevin Bugge, Charles Searby C, Darryl Nishimura, Kim A. Brogden, Joel N. Kline, Val Sheffield C and Michael J. Welsh. Proc Natl Acad Sci USA 2008; 105: 3380–3385.
[4] Annie Chiang P, Darryl Nishimura, Charles Searby, Khalil Elbedour, Rivka Carmi, Amanda Ferguson L, Jenifer Secrist, Terry Braun, Thomas Casavant, Edwin M. Stone and Val Sheffield C. Am J Hum Genet 2004; 75: 475–484.
[5] Maxence Nachury V, Alexander Loktev V, Qihong Zhang, Christopher Westlake J, Johan Peranen, Andreas Merdes, Diane Slusarski C, Richard Scheller H, Fernando Bazan J, Val Sheffield C, Peter Jackson K. Cell 2007; 129: 1201–1213.
[6] Bealesa PL, Elcioglub N, Woolfc AS, Parkerd D, Flintera FA. J Med Genet 1999; 36: 437-44
[7] Yip YL, Scheib H, Diemand AV, Gattiker A, Famiglietti LM, Gasteiger E and Bairoch A. Hum Mutat 2004; 23: 464–470.
[8] Yip YL, Famiglietti M, Gos A, Duek PD, David FP, Gateau A and Bairoch A. Hum Mutat 2008; 29: 361–366.
[9] Brigitte Boeckmann, Amos Bairoch, Rolf Apweiler, Marie-Claude Blatter, Anne Estreicher, Elisabeth Gasteiger, Maria Martin J, Karine Michoud, Claire O'Donovan, Isabelle Phan, Sandrine Pilbout, Michel Schneider. Nucleic Acids Res 2003; 31: 365–370.
[10] Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. Nucleic Acids Res 2000; 28: 235–242.
[11] Ng PC and Henikoff S. Nucleic Acids Res 2003; 31: 3812–3814.
[12] Ng PC and Henikoff S. Genome Res 2001; 11: 863–874.
[13] Ramensky V, Bork P and Sunyaev S. Nucleic Acids Res 2002; 30: 3894–3900.
[14] Capriotti E, Fariselli P and Casadio R. Nucleic Acids Res 2005; 33: 306–310.
[15] Bava KA, Gromiha MM, Uedaira H, Kitajima K and Sarai A. Nucleic Acids Res 2004; 32: 120–121.
[16] Gallivan JP, Dougherty DA. Proc Natl Acad Sci USA 1999; 96: 9459-64.
[17] Ramanathan K, Shanthi V, Rao Sethumadhavan. Int J Pharma Bio Sci 2010; 6: 1-15.
[18] Glaser F, Pupko T, Paz I, Bell RE, Bechor-Shental D, Martz E, Ben-Tal N. Bioinformatics 2003; 19:163-4.
[19] Kabsch W, Sander C. Biopolymers 1983; 22: 2577-2637.
[20] Heringa J, Argos P Proteins 1989; 37: 30-43.
[21] Tayubi IA and Sethumadhavan R Biochem 2010; 75: 912-918.
[22] Sudha Anand, Anand Anbarasu and Rao Sethumadhavan. In Silico Biol 2008; 8:261-73.

[23] Cavallo A and Martin AC. Bioinformatics 2005; 21: 1443–1450.
[24] Delarue M, Dumas P. Proc Natl Acad Sci USA 2004; 101: 6957–6962.
[25] Sharma S, Ding F, Nie H, Watson D, Unnithan A, Lopp J, Pozefsky D, Dokholyan NV. Bioinformatics 2006; 22: 2693–2694.
[26] Han JH, Kerrison N, Chothia C, Teichmann SA. Structure 2006; 14: 935–945.
[27] Varfolomeev SD, Uporov IV, Fedorov EV. Biochem 2002; 67: 1099–1108.
[28] Leach AR. Molecular Modeling: Principles and Applications. Pearson Education EMA, Sussex 2001.
[29] Chou KC, Carlacci L. Protein Eng 1991; 4: 661–667.
[30] Magyar C, Gromiha MM, Pujadas G, Tusnady GE, Simon. Nucleic Acids Res 2005; 33: 303–305.
[31] Tina KG, Bhadra R, Srinivasan N. PIC: Protein Interactions Calculator, Nucleic Acids Res 2007; 35: 473-476.
[32] Connolly M L. Sci 1983; 221: 709–713.
[33] Fabiola GF, Krishnaswamy S, Nagarajan V and Pattabi V. Acta Crystallogr D Biol Crystallogr 1997; 53: 316-320.
[34] Babu MM, Singh KS, Balaram P. J Mol Biol 2002; 322: 871-880.