Available Online at http://www.recentscientific.com

ISSN: 0976-3031

International Journal of Recent Scientific Research
Vol. 6, Issue, 6, pp.4644-4652, June, 2015

International Journal of Recent Scientific Research

RESEARCH ARTICLE

IDENTIFICATION AND STRUCTURAL COMPARISON OF DELETERIOUS MUTATIONS IN NSSNPS OF CYP1B1 GENE IN GLAUCOMA

Jincy Anna Jomy and Rao Sethumadhavan*

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

ARTICLE INFO

Article History:

Received 5th, May, 2015
Received in revised form 12th, May, 2015
Accepted 6th, June, 2015
Published online 28th, June, 2015

Key words:

Glaucoma, CYP1B1, AHR, protein docking, free energy.

ABSTRACT

We assessed the Single Nucleotide Polymorphisms (SNPs) that can change the expression and function in CYP1B1 gene using computational approaches. To study the connections among hereditary changes and phenotypic variation, distinctive computational methods like Sorting Intolerant from Tolerant (SIFT), Polymorphism Phenotyping (PolyPhen) and I-Mutant 3.0 are examined. There were 54 missense muatation; in this we watched 36 variations that were ruinous. We got 41 non-synonymous SNPs (nsSNPs) (72.92%) to be deleterious by SIFT. I- Mutant 3.0 were screening 52 (96.29%) deleterious mutants and 40 nsSNPs (74.07%) as from Polyphen-2. Cation-π interactions in protein structures are recognized and investigated the Arg, Lys interractions with π (Phe, Tyr or Trp) residues and their part in structural stability. Accordingly, modeling of these variations was implemented to comprehend the adjustment in their compliance concerning the native CYP1B1 by processing their root mean square deviation (RMSD). Those missense changes were because of loss of stability in their mutant structures of CYP1B1. The native and mutants were docked with the substrate 4f31 to clarify the substrate binding efficiencies of those hindering missense transformations. This was affirmed by figuring their total energies by utilizing GROMOS 96 force field and these changes were cross approved with computational projects. Here we presume that our work could recover the genes related missense mutations precisely as we cross validated some of our outcomes with tentatively demonstrated results by other research groups.

INTRODUCTION

The vast majority of the genetic issues are caused by biochemical inconsistencies (Goldstein, Polley, and Briggs, 2011). These are as often as possible created by point mutations indicated as single-nucleotide-polymorphisms (SNPs) (Zhang, Miteva, Wang, and Alexov, 2012). It can happen in any position of the genome which has consequences for the function of a gene. It can be found in both coding and non-coding regions (Casson, Chidlow, Wood, Crowston, and Goldberg, 2012). The non-synonymous SNPs (nsSNPs) cause changes in the amino acid residues (Liu, and Allingham, 2012). Glaucoma is a term describing a group of ocular disorders with multi-factorial etiology united by a clinically characteristic intraocular pressure-associated optic neuropathy (Gogate, Gilbert, and Zin, 2011). It is characterized by the progressive loss of retinal ganglion cells that is associated with a characteristic optic neuropathy and visual field loss. Primary open-angle glaucoma (POAG) and angle closure glaucoma (ACG) are the most prevalent forms of glaucoma and are the most common causes of glaucoma-related blindness worldwide (Lim *et al*., 2012). The disease manifests at birth or in the first year of life and usually leads to permanent vision impairment (Sheikh *et al*., 2014). The CYP1B1 gene was the first gene in which mutations were found to cause primary congenital glaucoma. It is located on chromosome 2p22–p21 (Achary and Nagarajam, 2008). Mutations in CYP1b1 have a deleterious effect on catalytic activity, leading to loss of function (Vasiliou and Gonzalez 2008). CYP1B1 protein is involved in the metabolism of steroids, retinol and retinal, arachidonate, and melatonin (Kaur, Mandal and Chakrabarti, 2011). CYP1B1 protein expresses in various human ocular tissues including cornea, ciliary body, iris, and retina (Gong *et al*., 2015). Sheikh *et al*., 2014 says CYP1B1 consists of two coding exons and encodes the cytochrome P450 superfamily, subfamily B, polypeptide 1, a 543 amino acids long protein. It is expressed in the trabecular meshwork and in the posterior segment of the eye. However, various in vitro and in-silico studies have demonstrated the pathogenic nature of the identified mutations. In spite of the fact that the accurate function of CYP1B1 protein in the eye is still hazy however as it is a mono-oxygenase, the accompanying situations may be normal for its part in the improvement of the eye.

MATERIALS AND METHODS

Datasets

The protein sequences and missense mutations of eye related genes were gotten from the Swissprot (Yip *et al*., 2008)/

Corresponding author:* **Rao Sethumadhavan
Bioinformatics Division, School of Bio Sciences and Technology, VIT University, Vellore, Tamil Nadu

Uniprot database and the 3D structure and its complex were acquired from Protein Data Bank (Westbrook *et al*., 2003) for in silico investigation and docking studies taking into account on detrimental point mutants.

Sequence Homology Based Method (Sift)

SIFT (accessible at http://sift.jcvi.org/) is a sequence homology-based instrument which is utilized to distinguish the destructive coding non-synonymous SNPs. SIFT accept that huge amino acids will be fixed in a protein family; accordingly, changes at all around saved positions have a penchant to be anticipated as unsafe (Ng and Henikoff, 2003). When we present the inquiries, SIFT takes the protein arrangement and uses different arrangement data for the expectation of endured and injurious substitutions for each position of question. In the wake of getting the question, SIFT looks for a comparative grouping given protein arrangement and it chooses the nearly related arrangements that can have comparable capacities. It gets the different arrangements of these chose successions and computes standardized probabilities for every single conceivable substitution at each position from the arrangement endured (Ng and Henikoff, 2001). In the event that the resilience record score is under 0.05, then it is anticipated to be pernicious and which is more prominent than 0.05 is thought to be nonpartisan.

Structure Homology Based Method (Polyphen)

Polyphen-2 (Ramensky *et al*., 2002) is a structural homology based programmed device which computes position-particular free numbers (PSIC) scores for each of the two variations and after that processes the PSIC score contrast between them. It is utilized for expectation of conceivable effect of an amino corrosive substitution on the structure and capacity of a human protein. Protein grouping is given as inquiry with a mutational position and two amino corrosive variations; it ascertains position-particular free tallies (PSIC) scores for each of the two variations and after that figures the PSIC scores contrast between them. In the event that the PSIC score distinction is higher, the useful effect of specific amino corrosive substitution is additionally higher will be liable to have. The parameters of PolyPhen-2 (accessible at http://genetics.bwh.harvard.edu/pph2/) server to figure the score are the succession based portrayal of the substitution site, report investigation of homologous groupings, and mapping of the substitution site to known protein 3D structures.

Support Vector Machine (I-Mutant 3.0)

I-Mutant 3.0 (accessible at http://folding.uib.es/cgi-receptacle/i-mutant3.0.cgi) is a support vector machine (SMV) based web server for the programmed forecast of protein solidness changes upon single-webpage transformations in light of Gibbs free energy (Capriotti *et al*., 2005). The apparatus was prepared on an information set got from ProTherm (Bava *et al*., 2004), which is right now the fundamental extensive database of exploratory information on protein changes. At the point when justified, I-Mutant 3.0 definitely forecast whether the protein transformation balances out or destabilizes the protein in 80% of the situations when the three-dimensional structure is known and 77% of the situations when just the protein grouping is accessible. The yield documents demonstrate the anticipated free vitality change esteem or sign (ΔΔG), which was computed from the unfolding Gibbs free vitality estimation of the transformed protein short the unfolding Gibbs free vitality estimation of the local protein (kcal mol^{-1}). Positive ΔΔG qualities predetermined that the changed protein has higher solidness and negative qualities show lower strength (Capriotti *et al*., 2005).

Mutant Modeling

To assess the basic soundness of local and mutant proteins, auxiliary investigation was performed by method for root mean square deviation (RMSD). The transformations were performed by utilizing SWISS PDB viewer and energy minimization for 3D structures was performed by NOMAD-Ref server (Lindahl *et al*., 2006) which utilizes Gromacs as default power field for energy minimization taking into account steepest descent, conjugate inclination and L-BFGS strategies (Delarue and Dumas, 2004). Here we have utilized the conjugate technique for advancing the 3D structures. Difference of the mutant structure from the native structure could be brought on by substitutions, erasures and insertions (Sharma *et al*., 2006) and the deviation between the two structures could modify the utilitarian movement (Han *et al*., 2006) concerning tying productivity of the inhibitors, which was assessed by their RMSD values.

Computation of Total Energy and Stabilizing Residues

To recognize the stability in the middle of native and mutant displayed structures, we discovered total energy which is processed by the GROMOS96 power field that is inserted in the SWISSPDB viewer. To foresee the energy of the atom as a component of its conformity, sub-atomic mechanics or power field routines use traditional sort models (Gromiha and Selvaraj, 2004). This permits expectation of equilibrium geometries, shift states and relative energies among conformers or among diverse atoms. Sub-atomic mechanics communicates the total energy as a total of Taylor arrangement extensions for the extends for each pair of reinforced molecules, and includes extra potential energy terms contributed by twisting, torsional energy, van der Walls energy, and electrostatics (Gromiha and Selvaraj, 2009). Consequently the total energy figuring could be measured as predictable requirement for comprehension the steadiness of protein particles with the assistance of Force field (Gromos96 and Gromacs). Performing energy minimization and reenacted toughening evacuates steric conflicts and serves to get the best consistent compliance (Varfolomeev *et al*., 2002). At long last, the total energy was assessed for native and mutants by the GROMOS power field. Furthermore, the total energy of the native structure was considered as a source of perspective point for looking at the total energy of mutant structures for steadiness investigation.

Furthermore, recognizing the settling buildups for both the native and mutant structures spoke to a huge parameter for comprehension their steadiness. Henceforth, we utilized the

server SRide (Magyar *et al*., 2005) to perceive the settling deposits in the native and mutant protein models. Settling deposits were processed utilizing parameters, for example, encompassing hydrophobicity, long-range request, adjustment focus, and protection score (Leach, 2001).

Computation of Cation–Π Interactions Energy

Cation-π interactions in protein structures are recognized and assessed by utilizing an energy based rule for selecting noteworthy side chain sets (Gallivan and Dougherty, 1999). These cation–π communications are gotten utilizing CaPTURE program. Cation-π collaborations are discovered to be basic among structures in the Protein Data Bank. The energies are registered for all the sets of cationic-aromatic amino acid deposits (Arg/Lys with Phe, Tyr and Trp) (Gromiha and Selvaraj, 2006). The aggregate Cation–π interaction energy (Ecat–π) has been separated into electrostatic (Ees) and van der Waals vitality (Evw) and was figured utilizing the system CaPTURE, which had actualized a subset of OPLS power field21 to figure the energies. The Ecat-π is the total of these two energies, i.e., electrostatic and the van der Waal's energy (Tayubi and Sethumadhavan, 2010).

Secondary Structure

We got the data about secondary structures DSSP (http://www.cmbi.ru.nl/dssp.html) (Wolfgang and Christian, 1983). The secondary structures were grouped into α-helix, β-strand, and irregular loop. The propensity of the amino acid residues to favor a specific conformation has been well documented. Such conformational preference is dependent not only on the amino acid alone but also on the native amino acid sequence. The secondary structure preferred of all o f the amino acids involved in all the above said types of C–H….π interactions were obtained using DSSP39 server. It is interesting to note that majority of the residues occupied in C–H….π interactions such as Phe, Trp, Tyr and His were chosen to be in β Strand.

Calculating the total number of intra molecular interactions using Pic Server

Protein Interactions Calculator (http://crick.mbu.iisc.ernet.in/~PIC) server is utilized for processing intra-sub-atomic communications for both native and mutant structures. PIC server accepts nuclear direction set of a protein structure in the standard Protein Data Bank (PDB) format. Interactions within a protein structure and collaborations between proteins in a get together are vital contemplations in comprehension sub-atomic premise of security and elements of proteins and their buildings. There are a few powerless and solid associations that render strength to a protein structure or a get together. It processes different 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 (Tina *et al*., 2007).

Identification of binding sites and computation of Atomic Contact Energy (ACE)

Here we utilized the system PatchDock for docking the native and mutants to process ACE by utilizing extra alternative of tying deposit parameter. The basic rule of this server is taking into account sub-atomic shape representation, surface patch coordinating in addition to sifting and scoring (Duhovny *et al*., 2002). It is gone for discovering docking changes that yield great sub-atomic shape complementarity. Such changes, when connected, prompt both wide interface territories and little measures of steric conflicts. A wide interface was guaranteed to incorporate a few coordinated nearby highlights of the docked atoms that have reciprocal qualities.

The PatchDock estimation partitions the connolly speck surface representation (Connolly, 1983) of the particles into sunken, raised and level patches. At that point, reciprocal patches were coordinated keeping in mind the end goal to produce applicant changes. Every competitor change was further assessed by a scoring capacity that considers both geometric fit and nuclear desolvation vitality (Sharma *et al*., 2006; Duhovny *et al*., 2005). At long last, a RMSD (root mean square deviation) alignment was connected to the competitor answers for toss repetitive arrangements. The fundamental explanation for Patch Dock's high proficiency was its quick transformational inquiry, which was driven by neighborhood highlight coordinating as opposed to animal power seeking of the six dimensional alter places. It further accelerates the computational handling time by using propelled information structures and spatial example detection methods, for example, geometric hashing and posture grouping.

At that point, the docked protein complex is given to the DFIRE server as information for figuring the coupling free energy (ΔG) scores (Zhang *et al*., 1997). It accepts another orientation state called the separation scaled, limited perfect gas reference (DFIRE) state. It is a separation ward structure-inferred potential grew so far and all utilized a reference express that can be portrayed as a deposit (molecule)-arrived at the midpoint of state. Likewise, the DFIRE-based all-molecule potential gives the most exact expectation of the dependable qualities of mutants taking into account information based all-iota possibilities (Yang and Zhou, 2008; Sreevishnupriya, *et al*., 2012).

RESULTS

Data collection

The SNPs and their related protein sequences for the CYP1B1 gene were gotten from the Swissprot (Yip *et al*., 2004) and the 3D structure (3PM0) has been obtained from PDB (Berman *et al*., 2000) for our computational examination.

SIFT

The conservation level of a particular position in a protein was managed by using a progression homology based instrument; SIFT (Ng and Henikoff, 2003).

Table1 List of functionally significant mutants predicted to be by I-Mutant 2.0, SIFT and PolyPhen.

S. No	Nucleotide change	AA change	Variants	SIFT	PolyPhen- 2	I-Mutant 3.0
1	L/P	77	L77P	0	1	-1.71
2	Y/N	81	Y81N	0	1	-1.47
3	A/P	115	A115P	0	1	-0.33
4	A/S	119	A119S	0.09	0	-0.73
5	M/R	132	M132R	0	0.948	-1.01
6	Q/H	144	Q144H	0.94	0.001	-0.69
7	Q/P	144	Q144P	0.04	0.976	-0.45
8	Q/R	144	Q144R	1	0.002	-0.11
9	R/W	145	R145W	0.02	0.931	-0.34
10	G/S	184	G184S	0.57	0.04	-1.27
11	A/P	189	A189P	0.12	0.88	-0.1
12	D/V	192	D192V	0	0.989	-0.19
13	P/L	193	P193L	0	1	-0.68
14	V/I	198	V198I	0.23	0.01	-0.85
15	N/S	203	N203S	0	1	-0.44
16	S/N	206	S206N	0.02	0.995	-0.54
17	S/I	215	S215I	0.01	0.907	0.06
18	E/K	229	E229K	0.01	0.656	-0.85
19	G/R	232	G232R	0.01	1	-0.48
20	S/R	239	S239R	0	1	0.08
21	R/L	266	R266L	0.31	0.057	-0.72
22	V/L	320	V320L	0.02	0.924	-1.24
23	A/F	330	A330F	0	1	-0.2
24	A/S	330	A330S	0	1	-0.61
25	L/F	345	L345F	0.15	0.265	-0.95
26	V/M	364	V364M	0	1	-1.07
27	G/W	365	G365W	0	1	-0.22
28	R/H	368	R368H	0	1	-1.31
29	D/N	374	D374N	0.01	0.997	-0.88
30	P/L	379	P379L	0.02	0.754	-0.55
31	E/K	387	E387K	0	1	-1.07
32	A/T	388	A388T	0.67	0.001	-0.84
33	R/C	390	R390C	0	1	-1.1
34	R/H	390	R390H	0	1	-1.65
35	R/S	390	R390S	0	1	-1.26
36	I/S	399	I399S	0	0.999	-2
37	V/F	409	V409F	0.01	0.519	-1.39
38	V/G	422	V422G	0	1	-2.21
39	N/Y	423	N423Y	0	1	0.03
40	L/V	432	L432V	0.29	0	-1.63
41	P/L	437	P437L	0	1	-0.57
42	D/H	441	D441H	0.01	0.977	-0.76
43	A/G	443	A443G	0.25	0	-1.4
44	R/Q	444	R444Q	0	1	-1.19
45	F/C	445	F445C	0	1	-1.89
46	D/E	449	D449E	0.18	0.001	-0.54
47	N/S	453	N453S	0.01	0.686	-0.6
48	G/D	466	G466D	0.02	1	-1.07
49	R/W	469	R469W	0	1	-0.59
50	E/G	499	E499G	0.02	0.947	-1.3
51	S/L	515	S515L	0.35	0.482	-0.14
52	V/A	518	V518A	0	0.64	-1.8
53	R/T	523	R523T	0	1	-0.94
54	D/G	530	D530G	0.07	0.008	-0.84

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

Protein preparations of 54 varieties were submitted self-rulingly to SIFT task to check its tolerance index. Higher the resistance record, lesser the valuable impact, a particular amino destructive substitution is inclined to have and the other path around. By this examination, all the 54 varieties showed noxious having the versatility rundown score some place around 0.01 and 0.00 as could be seen from Table 1.

Polyphen-2

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 0.519 to 1 (Table1).

I-Mutant 3.0

The nsSNPs are submitted to I-Mutant 3.0. Considering the differing Gibbs free essentialness estimation of changed and wild sort protein, (Table 1) of nsSNPs are found to be destabilize the protein (DDG<0 Kcal mol^{-1}).

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

S. No	Native Variants (AA change)	No. of SR	Stabilizing Residues
1	Native	2	PRO325, VAL326, PHE350, VAL351, PHE445
2	L77P	7	**PRO325, VAL326, VAL351**, ASN352, LEU438, **PHE445,** VAL447
3	Y81N	5	PHE323, **PRO325, PHE350**, LEU438, **PHE445**
4	A115P	4	VAL326, THR332, **PHE350, PHE445**
5	M132R	4	**PRO325, PHE350, PHE445**, VAL447
6	Q144P	2	VAL348, **PHE350**
7	R145W	5	VAL326, PRO329, **PHE350, PHE445**, VAL447
8	D192V	2	VAL348, **PHE350**
9	P193L	2	VAL348, **PHE350**
10	N203S	6	PHE323, **PRO325, PHE350**, LEU438, **PHE445,** VAL447
11	S206N	7	PHE323, **PRO325, PHE350**, VAL351, LEU438, **PHE445,** VAL447
12	E229K	3	**PRO325, PHE350,** VAL447
13	G232R	1	**PHE350**
14	V320L	3	**PHE350**, PHE445, VAL447
15	A330F	6	**PRO325,** VAL349, **PHE350,** LEU438, **PHE445,** VAL447
16	A330S	4	VAL326, THR332, **PHE350, PHE445**
17	V364M	5	PHE323, **PRO325, PHE350**, LEU438, **PHE445**
18	G365W	3	VAL326, **PHE350, PHE445**
19	R368H	3	VAL326, **PHE350, PHE445**
20	D374N	2	**PHE350, PHE445**
21	P379L	2	VAL348, **PHE350**
22	E387K	4	**PRO325, PHE350, PHE445**, VAL447
23	R390C	4	THR327, **PHE350,** THR439, VAL447
24	R390H	4	THR332, **PHE350, PHE445**, VAL447
25	R390S	4	VAL326, **PHE350,** ASN352, GLN353
26	I399S	4	**PRO325, PHE350, PHE445,** VAL447
27	V409F	3	**PRO325, PHE350, PHE445**
28	V422G	4	**PHE350,** GLY351, ASN352, **PHE445**
29	P437L	3	**PHE350, PHE445,** VAL447
30	D441H	1	**PHE350**
31	R444Q	5	VAL326, THR327, **PHE350, PHE445**, VAL447
32	F445C	2	VAL348, **PHE350**
33	N453S	8	PHE323, **PRO325,** THR327, **PHE350,** VAL351, LEU438, **PHE445,** VAL447
34	G466D	2	VAL348, **PHE350**
35	E499G	1	**PHE350**
36	V518A	2	**PHE350,** VAL351
37	R523T	3	**PRO325, PHE350, PHE445**

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

Rational consideration of Sift, Polyphen-2 And I-Mutant 3.0

We considered the 54 most potential hindering point changes for further course of examinations in light of the fact that they were generally discovered to be less steady, injurious, and harming by the I-Mutant 3.0, SIFT and Poly Phen-2 servers individually. The most commonly affected among the 3 computational tools has been taken for further studies i.e. 36 variants (Figure 1)

Computing RMSD and total energy by demonstrating native CYP1B1 and mutant structures

Mutations at indicated position were performed by SwissPDBViewer autonomously to get demonstrated structures. At that point, energy minimization is performed by NOMAD-Ref server (Lindahl *et al*., 2006) of both native structure and mutant demonstrated structures. To discover the deviation between the native and the mutant, the native structure was superimposed with all the energy refined mutant structures to get RMSD. The native structure of CYP1B1, Cytochrome P450 1B1 is 3PM0 which was taken from PDB. Table 2 shows the RMSD values for native structure with each mutant exhibited structure. It can be seen that, all the 36 mutants showed a RMSD regard between 0.45Å to 2.00Å. The higher the RMSD Value, the more is the deviation between the native and the mutant structure.

The end goal here is to discover the structural stability of native and mutants, the aggregate energy which bonds, angles, torsions, non-bonded and electrostatic constraints from GROMOS96 power field actualized in SwissPdb viewer to check their stability. It can be seen from Table 2 that the aggregate vitality for the native protein had -32049.725 kcal/mol while all the 36 mutants had the aggregate vitality higher than native protein. The superimposed structure of the native protein CYP1B1 with mutant type proteins (R444Q) is shown in Figure 2.

Computing stabilizing residues between native CYP1B1 and mutant modeled structures

We utilized the SRide server to see the settling stores of both the close-by structure and the mutant indicated structures (Table 2). This exhibits that the mutants, L77P, N203S, S206N, A330F and N453S were higher persevering than the native structure and Y81N, R145W, V364M and R444Q were showing similar number of stabilizing residues.

Intra-molecular interactions in TGFBI

We acknowledged the quality of protein structure by using the PIC server to perceive the amount of intra-sub-nuclear co operations for both local and mutant structures (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	1485	437	597	211	146	46	29	11	8
L77P	1829	435	700	351	235	55	26	12	15
Y81N	1816	434	701	343	232	56	24	12	14
A115P	1848	442	718	344	239	55	24	11	15
M132R	1854	443	709	366	230	53	28	11	14
Q144P	1557	443	642	222	151	49	29	11	10
R145W	1865	449	718	361	237	51	26	10	13
D192V	1534	441	625	223	147	49	29	10	10
P193L	1534	441	625	223	147	49	29	10	10
N203S	1862	432	712	364	246	57	24	12	15
S206N	1858	440	707	365	240	53	29	11	13
E229K	1842	438	700	358	240	53	27	12	14
G232R	1539	440	623	227	149	51	29	10	10
V320L	1877	445	716	370	243	54	26	9	14
A330F	1876	455	715	357	242	52	29	12	14
A330S	1857	438	706	376	228	55	28	11	15
V364M	1835	434	701	358	236	55	23	13	15
G365W	1890	453	713	378	236	54	29	15	12
R368H	1828	429	702	357	233	56	24	11	16
D374N	1851	444	708	363	234	53	26	10	13
P379L	1554	437	643	225	150	49	29	11	10
E387K	1832	440	716	355	216	51	26	11	17
R390C	1855	437	702	372	239	52	27	11	15
R390H	1870	450	718	351	246	52	30	11	12
R390S	1904	443	716	384	254	52	28	11	16
I399S	1830	418	714	353	237	56	23	12	17
V409F	1834	454	695	349	226	51	32	14	13
V422G	1829	424	711	347	243	56	23	11	14
P437L	1867	455	713	350	241	52	31	11	14
D441H	1547	441	637	224	147	48	29	11	10
R444Q	1831	438	708	346	237	51	26	11	14
F445C	1533	433	630	226	147	49	27	11	10
N453S	1850	435	719	347	243	55	25	11	15
G466D	1539	439	625	227	147	51	29	11	10
E499G	1834	445	702	366	223	51	26	9	12
V518A	1879	452	708	377	238	53	26	11	14
R523T	1771	424	696	344	204	53	23	12	15

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

There were a couple of weak and strong intra-nuclear affiliations that render strength to a protein structure. Hence these intra-sub-nuclear coordinated efforts were figured by PIC server to further substantiate the quality of protein structure. In perspective of this observation, we found that a total number of 1485 intra-nuclear joint efforts were gotten in the nearby structure of TGFBI. On the other hand, 19 mutant structures of TGFBI constructed the intra-sub-nuclear joint efforts between the extent of 415 to 441.

Cation–Π Interactions

The Cation– π interaction energy of both native and mutant were investigated. There were four sets of cation-π interactions (Arginine- Phenylalanine, Arginine- Tyrosine, Arginine-Tryptophan and Lysine- Tyrosine) in native are -5.47, -3.34, -6.85 and -3.85 respectively (Table 4). Then again mutants demonstrate the same cation-π interactions. So this shows three mutants secured marginally solid cation-π connection than Native. The composition of native cation-π interactions are demonstrated in Figure 3.

Secondary Structure

Here we have calculated the happening of cation-π forming residues in secondary structure particularly in native structure. We found that the cation-π forms Alpha helix (H), Bend (S) and Helix (G) which is shown in Table 5.

Analyzing the binding efficiency for native KIF11 and mutant

Here we performed 36 missense changes in A chain of local 3PM0 and communicating protein 4F31 by swisspdb viewer uninhibitedly and vitality minimization was executed for the whole complex (both local and mutant complex) by GROMACS (Nomad-ref) trailed by mimicked fortifying to get the overhauled structures using a discrete sub-atomic development approach (ifold). We utilized PatchDock to dock native 3PM0 and mutant structures with 4F31 furthermore; we utilized DFire, for finding the protein suitability free vitality hotspot for the docked complex recovered from PatchDock. We utilized this server for the missense change examination concerning discovering the free centrality wellspring of both local and mutants of 4F31 (Table 5). In this examination, we found that the coupling tying free energy for 4F31 with native 3PM0 protein was discovered to be -1288.06kcal/mol, which have to some degree higher tying regular inclination separated from the mutants. This examination depicts that area 3PM0 showed higher tying proclivity with 4F31.

Table 4 Average cation-π interaction energy

Variants	R-F(-Kcal/mol)	R-Y(-Kcal/mol)	R-W(-Kcal/mol)	K-Y(-Kcal/mol)
Native	R158-F155(-5.47)	R80-Y81(-3.34)	R390-W434(-6.85)	K142-Y137(-3.85)
L77P	R158-F155(-5.47)	R80-Y81 (-3.34)	R390-W434 (-6.85)	K142-Y137 (-3.85)
Y81N	R158-F155(-5.47)	R80-Y81 (-3.34)	R390-W434 (-6.85)	K142-Y137 (-3.85)
A115P	R158-F155(-5.47)	R80-Y81 (-3.34)	R390-W434 (-6.85)	K142-Y137 (-3.85)
M132R	R158-F155(-5.47)	R80-Y81 (-3.34)	R390-W434 (-6.85)	K142-Y137 (-3.85)
Q144P	R158-F155(-5.47)	R80-Y81 (-3.34)	R390-W434 (-6.85)	K142-Y137 (-3.85)
R145W	R158-F155(-5.47)	R80-Y81 (-3.34)	R390-W434 (-6.85)	K142-Y137 (-3.85)
D192V	R158-F155(-5.47)	R80-Y81 (-3.34)	R390-W434 (-6.85)	K142-Y137 (-3.85)
P193L	R158-F155(-5.47)	R80-Y81 (-3.34)	R390-W434 (-6.85)	K142-Y137 (-3.85)
N203S	R158-F155(-5.47)	R80-Y81 (-3.34)	R390-W434 (-6.85)	K142-Y137 (-3.85)
S206N	R158-F155(-5.47)	R80-Y81 (-3.34)	R390-W434 (-6.85)	K142-Y137 (-3.85)
E229K	R158-F155(-5.47)	R80-Y81 (-3.34)	R390-W434 (-6.85)	K142-Y137 (-3.85)
G232R	R158-F155(-5.47)	R80-Y81 (-3.34)	R390-W434 (-6.85)	K142-Y137 (-3.85)
V320L	R158-F155(-5.47)	R80-Y81 (-3.34)	R390-W434 (-6.85)	K142-Y137 (-3.85)
A330F	R158-F155(-5.47)	R80-Y81 (-3.34)	R390-W434 (-6.85)	K142-Y137 (-3.85)
A330S	R158-F155(-5.47)	R80-Y81 (-3.34)	R390-W434 (-6.85)	K142-Y137 (-3.85)
V364M	R158-F155(-5.47)	R80-Y81 (-3.34)	R390-W434 (-6.85)	K142-Y137 (-3.85)
G365W	R158-F155(-5.47)	R80-Y81 (-3.34)	R390-W434 (-6.85)	K142-Y137 (-3.85)
R368H	R158-F155(-5.47)	R80-Y81 (-3.34)	R390-W434 (-6.85)	K142-Y137 (-3.85)
D374N	R158-F155(-5.47)	R80-Y81 (-3.34)	R390-W434 (-6.85)	K142-Y137 (-3.85)
P379L	R158-F155(-5.47)	R80-Y81 (-3.34)	R390-W434 (-6.85)	K142-Y137 (-3.85)
E387K	R158-F155(-5.47)	R80-Y81 (-3.34)	R390-W434 (-6.85)	K142-Y137 (-3.85)
R390C	R158-F155(-5.47)	R80-Y81 (-3.34)	R390-W434 (-6.85)	K142-Y137 (-3.85)
R390H	R158-F155(-5.47)	R80-Y81 (-3.34)	R390-W434 (-6.85)	K142-Y137 (-3.85)
R390S	R158-F155(-5.47)	R80-Y81 (-3.34)	R390-W434 (-6.85)	K142-Y137 (-3.85)
I399S	R158-F155(-5.47)	R80-Y81 (-3.34)	R390-W434 (-6.85)	K142-Y137 (-3.85)
V409F	R158-F155(-5.47)	R80-Y81 (-3.34)	R390-W434 (-6.85)	K142-Y137 (-3.85)
V422G	R158-F155(-5.47)	R80-Y81 (-3.34)	R390-W434 (-6.85)	K142-Y137 (-3.85)
P437L	R158-F155(-5.47)	R80-Y81 (-3.34)	R390-W434 (-6.85)	K142-Y137 (-3.85)
D441H	R158-F155(-5.47)	R80-Y81 (-3.34)	R390-W434 (-6.85)	K142-Y137 (-3.85)
R444Q	R158-F155(-5.47)	R80-Y81 (-3.34)	R390-W434 (-6.85)	K142-Y137 (-3.85)
F445C	R158-F155(-5.47)	R80-Y81 (-3.34)	R390-W434 (-6.85)	K142-Y137 (-3.85)
N453S	R158-F155(-5.47)	R80-Y81 (-3.34)	R390-W434 (-6.85)	K142-Y137 (-3.85)
G466D	R158-F155(-5.47)	R80-Y81 (-3.34)	R390-W434 (-6.85)	K142-Y137 (-3.85)
E499G	R158-F155(-5.47)	R80-Y81 (-3.34)	R390-W434 (-6.85)	K142-Y137 (-3.85)
V518A	R158-F155(-5.47)	R80-Y81 (-3.34)	R390-W434 (-6.85)	K142-Y137 (-3.85)
R523T	R158-F155(-5.47)	R80-Y81 (-3.34)	R390-W434 (-6.85)	K142-Y137 (-3.85)

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

Table 5 Secondary structure Preferences of KIF11

PDB ID	Cat-residue	2° str	ASA	π-residue	2° str	ASA	D_{seq}
3PM0	R158	S	123	F155	S	16	3
	R80	H	201	Y81	H	125	1
	R390	H	0	W434	G	2	44
	K142	H	49	Y137	S	50	5

Notes: 20 Str: Secondary structure, H- Alpha helix, S- Strand, G- Helix, ASA: Accessible Surface Area, Dseq: Sequence distance of separation between cationic and π residues

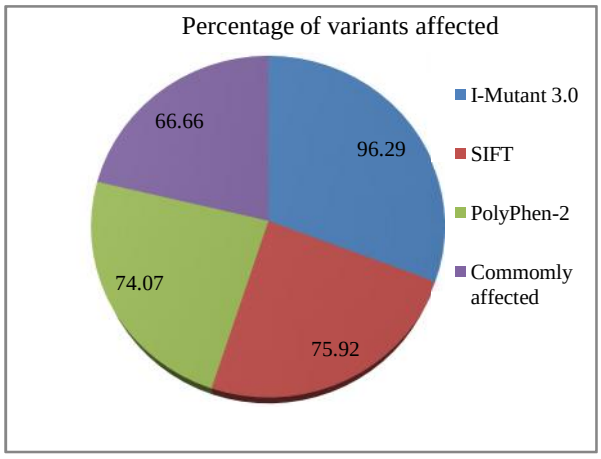


Figure 1 List of functionally significant mutations.

In this way, the lesser tying free energies might likely be an immediate consequence of loss of intermolecular non-covalent joint efforts.

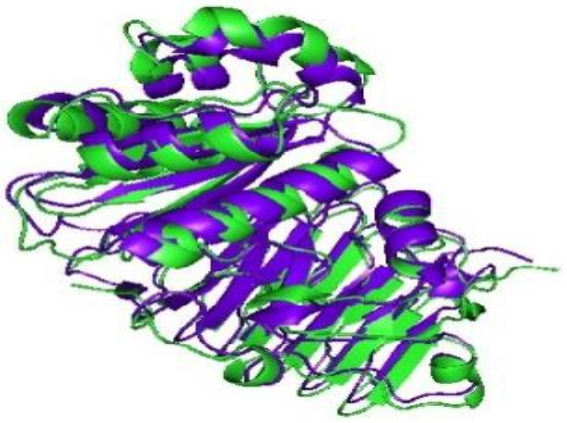


Figure 2 Superimposed structure of native CYP1B1 protein (green) with mutant R444Q (violet) structure showing RMSD of 2.04Å

This examination plainly depicted that native complex had high intermolecular non covalent correspondences than mutant structure.

Table 6 Atomic Contact Energy and Free Energy for the native and mutants

S.No	Variants	Atomic Contact Energy(ACE)	Free Energy (Kcal/mol)
1	Native	-406.94	-1288.06
2	L77P	-407.33	-1314.67
3	Y81N	-482.06	-1306.64
4	A115P	-323.21	-1313.27
5	M132R	-6.18	-1306.19
6	Q144P	-268.58	-1284.14
7	R145W	-125.32	-1319.08
8	D192V	-338.78	-1302.1
9	P193L	-338.78	-1302.1
10	N203S	-254.8	-1316.42
11	S206N	-339.37	-1308.69
12	E229K	-174.28	-1319.95
13	G232R	-180.07	-1287.57
14	V320L	-322.63	-1314.56
15	A330F	-172.91	-1322.48
16	A330S	-470.97	-1310.3
17	V364M	-233.11	-1308.65
18	G365W	-219.56	-1326.4
19	R368H	-452.91	-1314.58
20	D374N	-420.69	-1321.4
21	P379L	-421.05	-1279.07
22	E387K	-203.43	-1298.95
23	R390C	-160.86	-1308.72
24	R390H	-84.87	-1308.72
25	R390S	-233.39	-1310.07
26	I399S	-143.32	-1314.67
27	V409F	-470.76	-1316.88
28	V422G	-262.49	-1309.19
29	P437L	-391.02	-1315.23
30	D441H	-323.26	-1284.22
31	R444Q	-387.07	-1323.03
32	F445C	-257.52	-1284.72
33	N453S	-392.47	-1319
34	G466D	-309.14	-1282.48
35	E499G	-245.06	-1315.05
36	V518A	-5.2	-1317.95
37	R523T	-497.54	-1304.23

Atomic contact energy and Free energy

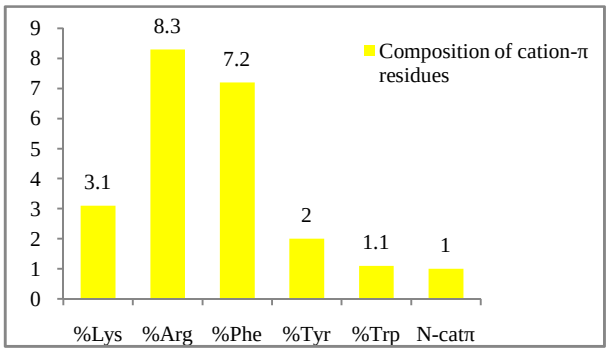
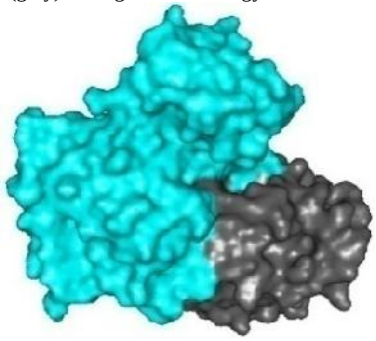


Figure 4 Docked complex of Native CYP1B1 (cyan) and AHR (grey) having the Free energy of -1288.06



CONCLUSION

In this study, the CYP1B1 protein which has been discovered to be connected with glaucoma was researched by computational investigation for harmful missense transformations. Out of the 54 mutants in the alpha chain of CYP1B1 protein, 36 was discovered to be harming by PolyPhen server, SIFT and I-Mutant 3.0. Docking examination between native and mutants with the partner CYP1B1 and AHR furthermore, structures created Atomic Contact Energy scores -5.2 and -497.54 separately Therefore the outcomes show that our methodology effectively permitted us to (1) think about computationally as a suitable convention for missense change (point transformation/single amino corrosive polymorphism) investigation before wet lab experimentation and (2) gave an ideal way to further clinical and test studies to portray CYP1B1 mutants inside and out.

References

A. R. Leach. 2001. Molecular Modeling: Principles and Applications. Vol. 2, Pearson Education EMA, Sussex

Achary, M.S. and Nagarajam, H.A. 2008. Comparative docking studies of CYP1b1 and its PCG-associated mutant forms. J Biosci. 33(5): 699-713.

Bava, K.A., Gromiha, M.M., Uedaira, H., Kitajima, K. and Sarai, A. 2004. Protherm, Version 4.0: thermodynamic Database for proteins and mutants. Nucleic Acids Res. 32: 120–121.

Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N. and Bourne, P.E. 2000. The Protein Data Bank. Nucleic Acids Res. 28: 235–242

Casson, R.J, Chidlow, G., Wood, J.P., Crowston, J.G., Goldberg, I. (2012). Definition of glaucoma: clinical and experimental concepts. Clin Experiment Ophthalmol., 40(4):341-9.

Connolly, M.L.1983. Analytical molecular surface calculation. J. Appl. Cryst. 16: 548-558.

Delarue, M. and Dumas, P. 2004. On The Use Of Low-Frequency Normal Modes To Enforce Collective Movements In Refining Macromolecular Structural Models. Proc Natl Acad Sci USA. 101(8), pp. 6957–6962.

Dina, S., Andrea, R., Agustin Avila-Sakar, A., Seung, J.K., Velázquez-Muriel, J., Pavel, S., Hong, L., Kristin, A. K., Maofu, L., Ho Min, K., Solmaz, S., Volker, D., Arvind, R., Jaume, P., David, A.A., Yifan, C. and Andrej, S. 2012. A method for integrative structure determination of protein-protein complexes. Bioinformatics. 28 (24): 3282–3289.

Gallivan, J.P. and Dougherty, D.A. 1999. Cation-pi interactions in structural biology, Proc Natl Acad Sci U S A. 96(17): 9459-64.

Gogate, P., Gilbert, C. and Zin, A. 2011. Severe Visual Impairment and Blindness in Infants: Causes and Opportunities for Control. Middle East Afr J Ophthalmol. 18(2): 109–114.

Goldstein, B. A., Polley E. C. and Briggs, F.B.S. (2011). Random Forests for Genetic Association Studies. Stat Appl Genet Mol Biol. 10(1), 32.

Gong, B., Chao, Q., Xiulan, L., Yi, S., Ying L., Zhou Y., Shuai, P., Yin Y., Xiaoqi L., Dingding, Z. and Zhenglin Y. 2015. Mutation spectrum of CYP1B1 in Chinese patients with primary open-angle glaucoma. Br J Ophthalmol. 99: 425-430.

Gromiha, M.M. and Selvaraj S. 2006. Salient Features of Residue Contacts in Protein Structures for Determining their Folding and Stability. In "Protein Folding: New Research", Nova Science Publishers, New York, USA. 249-265.

Gromiha, M.M. and Selvaraj, S. 2004. Inter-residue interactions in protein folding and stability. Prog Biophys Mol Biol. 86(2): 235-77.

Gromiha, M.M. and Selvaraj, S. 2009. Proteins: computational analysis of structure, function and stability. Wiley Encyclopedia on Chemical Biology 4, 174-182.

Han, J.H., Kerrison, N., Chothia, C. and Teichmann, S.A. Divergence of interdomain geometry in two-domain proteins. Structure. 14(5): 935–945.

Kabsch, W. and Sander, C. 1983. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features, Biopolymers, 22: 2577-2637.

Kaur, K., Anil, K.M. and Chakrabarti, S., 2011. Primary Congenital Glaucoma and the Involvement of CYP1B1. Middle East Afr J Ophthalmol. 18(1): 7–16.

Lindahl, E., Azuara, C., Koehl, P. and Delarue, M. 2006. NOMAD-Ref: Visualization, Deformation And Refinement Of Macromolecular Structures Based On All-Atom Normal Mode Analysis. Nucleic Acids Res. 34: 52–56.

Liu, Y. and Allingham R.R. 2011. Molecular Genetics in Glaucoma. Exp Eye Res., 93(4): 331–339.

Magyar, C., Gromiha, M.M, Pujadas, G., Tusnády, G.E. and Simon, I. 2005. SRide: a server for identifying stabilizing residues in proteins, Nucleic Acids Res. 33: 303–305

Ng, P.C. and Henikoff, S. 2003. SIFT: Predicting Amino Acid Changes That Affect Protein Function. Nucleic acids res. 31(13): 3812–3814.

Ramensky, V., Bork P. and Sunyaev, S. 2002. Human non-synonymous SNPs: server and survey. Nucleic Acids Res. 30(7): 3894–3900

Shakeel, A.S., Waryah, A.M., Narsani, A.K., Hina, S., Imtiaz, A.G., Khairuddin, S., Muhammad, Q., Azam, I.M., Pitambar K. and Shaikh, N. 2014. Mutational spectrum of the CYP1B1 gene in Pakistani patients with primary congenital glaucoma: Novel variants and genotype-phenotype correlations. Mol Vis. 20: 991–1001.

Sharma, S., Ding, F., Nie, H., Watson, D., Unnithan, A., Lopp, J., Pozefsky, D. and Dokholyan, N.V. 2006. iFold: a platform for interactive folding simulation of proteins. Bioinformatics. 22(21), pp. 2693–2694

Sing-Hui, L., Khanh-Nhat, T.V., Tammy, L., Yanovitch, S., Freedman, F., Thomas, K. Whitney, C., Caldwell, P., Ajay, R., Ravikanth, M., Erica, B.N., Steve, R. and Terri L.Y. 2013. CYP1B1, MYOC, and LTBP2 Mutations in Primary Congenital Glaucoma Patients in the United States. Am J Ophthalmol. 155(3): 508–517

Sreevishnupriya, K., Chandrasekaran, P., Senthilkumar, A., Sethumadhavan, R., Shanthi, V., Daisy, P., Nisha, J., Ramanathan, K. and Rajasekaran, R. 2012. Computational analysis of deleterious missense mutations in aspartoacylase that cause Canavan's disease, Science China Life sciences, 55: 1190-1119.

Tayubi, I.A. and Sethumadhavan, R. 2010. Nature of cation-pi interactions and their role in structural stability of immunoglobulin proteins. Biochemistry, 75: 912-8.

Tayubi, I.A., Sethumadhavan, R. (2011). Theoretical understanding of C–H….Π interactions and their distribution in immunoglobulin proteins- insilco geometrical approach, Int J Pharm Pharm Sci, 3: 212- 218

Tina, K.G., Bhadra, R. and Srinivasan, N. 2007. PIC: Protein Interactions Calculator, Nucleic Acid Research, 35: 473-6.

Varfolomeev, S.D., Uporov, I.V. and Fedorov, E.V. 2002. Bioinformatics and Molecular Modeling in Chemical Enzymology. Active Sites of Hydrolases. Biochem. (Mosc). 67: 1099–1108.

Vasiliou, V. and Gonzalez, F.J. 2008. Role of CYP1B1 in glaucoma. Annu Rev Pharmacol Toxicol. 48: 333-58.

Yang, Y. and Zhou, Y. 2008. Ab initio folding of terminal segments with secondary structures reveals the fine difference between two closely related all-atom statistical energy functions. Prot Sci. 17 : 1212-1219.

Yip, Y.L, Scheib, H., Diemand, A.V, Gattiker, A., Famiglietti LM, Gasteiger E and Bairoch, A. 2004. The Swiss-Prot Variant Page and the Modsnp Database: A Resource For Sequence And Structure Information On Human Protein Variants. Hum Mutat. 23(5): 464–470.

Zhang, B., Gallegos, M., Puoti, A., Durkin, E., Fields, S., Kimble, J. and Wickens, M.P. 1997. A conserved RNA-binding protein that regulates sexual fates in the C. elegans hermaphrodite germ line. Nature. 390(6659):477-84.

Zhang, Z., Miteva, M. A., Wang, L. and Alexov, E. (2012) Analyzing Effects of Naturally Occurring Missense Mutations. Comput Math Methods Med., 2012: 805827.

How to cite this article:

Jincy Anna Jomy and Rao Sethumadhavan., Identification And Structural Comparison Of Deleterious Mutations In Nssnps Of Cyp1b1 Gene In Glaucoma. *International Journal of Recent Scientific Research Vol. 6, Issue, 6, pp.4644-4652, June, 2015*

****