Computers in Biology and Medicine 107 (2019) 161–171

Contents lists available at ScienceDirect

Computers in Biology and Medicine

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

Understanding the structure-function relationship of *HPRT1* missense mutations in association with Lesch–Nyhan disease and HPRT1-related gout by in silico mutational analysis

Ashish Kumar Agrahari[a], M. Krishna Priya[a], Medapalli Praveen Kumar[a], Iftikhar Aslam Tayubi[b], R. Siva[a], B. Prabhu Christopher[c], C. George Priya Doss[a,*], Hatem Zayed[d,**]

[a] Department of Integrative Biology, School of Biosciences and Technology, VIT, Vellore, Tamil Nadu 632014, India
[b] Faculty of Computing and Information Technology, King Abdulaziz University, Rabigh, 21911, Saudi Arabia
[c] VIT-BS, VIT, Vellore, Tamil Nadu, 632014, India
[d] Department of Biomedical Sciences, College of Health and Sciences, Qatar University, Doha, Qatar

ARTICLE INFO

Keywords:
HPRT1
Lesch–nyhan disease
Non-synonymous SNP
Molecular dynamics simulation

ABSTRACT

The nucleotide salvage pathway is used to recycle degraded nucleotides (purines and pyrimidines); one of the enzymes that helps to recycle purines is hypoxanthine guanine phosphoribosyl transferase 1 (HGPRT1). Therefore, defects in this enzyme lead to the accumulation of DNA and nucleotide lesions and hence replication errors and genetic disorders. Missense mutations in hypoxanthine phosphoribosyl transferase 1 (*HPRT1*) are associated with deficiencies such as Lesch–Nyhan disease and chronic gout, which have manifestations such as arthritis, neurodegeneration, and cognitive disorders. In the present study, we collected 88 non-synonymous single nucleotide polymorphisms (nsSNPs) from the UniProt, dbSNP, ExAC, and ClinVar databases. We used a series of sequence-based and structure-based in silico tools to prioritize and characterize the most pathogenic and stabilizing or destabilizing nsSNPs. Moreover, to obtain the structural impact of the pathogenic mutations, we mapped the mutations to the crystal structure of the HPRT protein. We further subjected these mutant proteins to a 50 ns molecular dynamics simulation (MDS). The MDS trajectory showed that all mutant proteins altered the structural conformation and dynamic behavior of the HPRT protein and corroborated its association with LND and gout. This study provides essential information regarding the use of HPRT protein mutants as potential targets for therapeutic development.

1. Introduction

Lesch–Nyhan disease (LND) is a rare X-linked neurogenetic metabolic disorder caused by mutations in the hypoxanthine phosphoribosyl transferase 1 (*HPRT1*) gene, which encodes HGPRT, an enzyme that salvages purines from degraded DNA and reintroduces purines in the purine salvage pathway [1–3]. These mutations can be induced by UV exposure when the DNA repair machinery responsible for UV-induced damage fails, mainly due to mutations in the proteins functioning in nucleotide excision repair (NER) and mismatch repair (MMR). In fact, an HPRT mutation frequency assay was developed based on *HPRT1* mutations as a measure of UV-induced DNA lesions to diagnose the damage response [4–7]. Some mutations that do not associate with LND (juvenile gout disease) give rise to HPRT-linked gout (adult gout disease) [8,9]. LND results from the partial or complete loss of HGPRT enzyme activity [1]. The characteristic features of LND include hyperuricemia, early-onset hypotonia with subsequent onset of a primarily dystonic movement disorder, dysarthric speech, intellectual disability, and compulsive self-injury with self-mutilation along with an extended cognitive and behavioral phenotype [1]. Partial HGPRT enzyme deficiency causes excess uric acid production and a variable scale of neurological manifestations leading to Lesch–Nyhan variants (LNVs) [10,11]. The mildest variant was characterized by isolated excess uric acid production, no clinically explicit neurologic or behavioral defects, and most often HGPRT-related hyperuricemia (HRH) [12,13]. The intermediate phenotype manifests with a range of varying degrees of neurological and behavioral defects, termed HGPRT-related neurological dysfunction (HND) [14]. HND patients suffer from excess uric acid

* Corresponding author. Department of Integrative Biology, School of Biosciences and Technology, VIT, Vellore, India.
** Corresponding author.
E-mail addresses: georgepriyadoss@vit.ac.in (C. George Priya Doss), hatem.zayed@qu.edu.qa (H. Zayed).

https://doi.org/10.1016/j.compbiomed.2019.02.014
Received 30 November 2018; Received in revised form 13 February 2019; Accepted 19 February 2019
0010-4825/ © 2019 Elsevier Ltd. All rights reserved.

production accompanied by some neurological and behavioral abnormalities; however, they do not manifest self-injury as observed in classic LND [14]. The genetic etiology of LND involves a mutation in the *HPRT1* gene, which is located on the long arm of the X-chromosome (Xq26.1) and is constituted of 9 exons and 8 introns [11]. The gene is located on the X-chromosome. Therefore, the disorder mainly affects males rather than females [15]. To date, over 615 mutations (http://www.lesch-nyhan.org/en/research/mutations-database) have been reported in the *HPRT1* gene [11]. Among these, 242 are missense mutations, of which 131 are reported to be associated with LND. GLUT9 SNPs also influence the uric acid concentration in patients with Lesch-Nyhan disease [16]. Few clinical studies have indicated that the mutations L41P [17], G70E [17–20], G71R [21,22], I132T [23] and M143K [24] are associated with HPRT deficiency and linked to classical LND. The predominant phenotype in most patients with Lesch-Nyhan syndrome is a decreased enzyme concentration instead of an abnormality in enzyme function [25]. However, it has been reported that the G71R substitution showed no detectable catalytic activity of the HGPRT enzyme and cathodal migration in the native PAGE. The probability of a change in beta-turn formation as a consequence of the G71R substitution was predicted using Chou-Fasman secondary structure analysis [21,22]. G70E presented HPRT deficiency with classical Lesch–Nyhan symptoms [17–20]. L41P [17], I132T [23], and M143K [24] substitutions also resulted in no detectable catalytic activity of the HGPRT enzyme. In the current study, we were mainly interested in filtering out the most pathogenic (disease-causing) mutations from the pool of mutations available in public databases.

Single nucleotide polymorphisms (SNPs) are the most common genetic substitutions in humans and are defined by an allele frequency of 1% or more [26,27]. Based on whether the amino acid of the encoded protein changes or not, SNPs can be classified as non-synonymous SNPs (nsSNPs) or synonymous (sSNPs), respectively. As only approximately 1.5% of the human genome codes for proteins, most of the SNPs arise outside coding regions. These SNPs or variations may have an impact on gene expression and regulation by altering the regulatory regions and disturbing transcription factor-binding sites. NsSNPs are of particular interest because they result in altered amino acid sequences of encoded proteins, and consequently, residue substitutions may result in modified structures and/or functions of proteins [28–30]. SNPs are of vital clinical importance due to the possible association of these SNPs with different traits and genetic diseases. Although nsSNPs can affect protein structure and function, not every nsSNP gives rise to considerable changes in protein structure. Therefore, in silico prediction tools of different algorithms are used to predict the pathogenic and destabilizing nature of nsSNP etiology [31–35].

In the present study, we utilized a list of sequence-based, structure-based algorithms along with stability prediction tools for screening and identifying the most deleterious and pathogenic nsSNPs. To investigate the potential effects of nsSNPs on HPRT protein structure and function, we examined the structural and dynamic behavior of wild-type and mutant HPRT using molecular dynamics simulation (MDS). Several previous studies have successfully illustrated the significance of MDS in combination with in silico SNP prediction tools for understanding the molecular impact of nsSNPs on protein structure and function [33,34,36–50]. The nsSNPs that were filtered as the most pathogenic and reported to be clinically harmful were selected for further study. Moreover, pathogenic nsSNPs were mapped to the HPRT structure to understand their effects at the structural level. Finally, we performed a 50ns MDS for wild-type and mutant HPRT to gain insights into noteworthy molecular dynamics behavior of the proteins.

2. Materials and methods

2.1. Data collection

Information about the nsSNPs was collected from the UniProt database (https://www.uniprot.org) [51] with the UniProt ID: P00492, dbSNP database [52] (https://www.ncbi.nlm.nih.gov/snp) ExAC (http://exac.broadinstitute.org/) and ClinVar (https://www.ncbi.nlm.nih.gov/clinvar) for LND and gout. The FASTA sequence for HPRT was downloaded from the UniProt database and used for further analysis. The three-dimensional crystal structure of the HPRT protein was retrieved from the PDB database (PDB ID: 1BZY) (https://www.rcsb.org/) [53].

2.2. In silico screening of deleterious nsSNPs

Since not every *HPRT1* mutation is damaging and not every LND disease is caused by an *HPRT1* mutation [54], we used a series of in silico tools to prioritize the most pathogenic nsSNPs from the pool of nsSNP information available in public databases. Initially, we utilized sequence-based tools (SIFT [55], SNAP [56], FATHMM [57]) and structure-based PolyPhen 2 [58] to predict pathogenic nsSNPs from non-pathogenic or neutral nsSNPs in the coding region. SIFT (sorting intolerant from tolerant) is one of the standard tools used to characterize missense mutations. It predicts whether an amino acid substitution alters functionality by designating the substitution as neutral or deleterious using sequence homology. SIFT provides a normalized probability score; if the score is less than 0.05, then the amino acid substitution is deemed deleterious, and if the score is greater than 0.05, then the substitution is deemed neutral [55]. SNAP is a program that uses neural networks for the prediction of missense variants and based on the structural feature of proteins; it segregates the variants as neutral and non-neutral. Structural features such as secondary structure and solvent accessibility are considered to characterize a position in the protein. Three outputs are given in SNAP for each amino acid substitution: reliability index from 0 to 9, binary prediction, and expected accuracy [56].

FATHMM (functional analysis through hidden Markov models) is an algorithm that combines sequence conservation to predict the impact of a missense mutation on the structural and phenotypic characteristics of a protein. The FATHMM server input is the UniProt ID with amino acid variation, prediction algorithm, and phenotypic associations are provided [57]. PolyPhen-2 is a tool that makes use of sequence and structure-based information to predict the effect of amino acid substitution on the function of a protein. It involves a comparison between the wild-type allele and the corresponding mutant or disease-causing allele. PolyPhen-2 calculates the position-specific independent count score, from which the substitution can be classified as benign (0, 0.02), possibly damaging (0.02–0.85) or most likely damaging (0.85–1) [58].

Furthermore, we predicted the stabilizing or destabilizing effects of mutations using I-mutant3.0 [59] and iStable [60]. I-mutant3.0 uses a support vector machine (SVM)-based algorithm for the prediction of the impact of an amino acid substitution on protein stability. It calculates the change in free energy (DDG) upon mutation. If the DDG score is less than -0.5, then the mutation primarily destabilizes the protein, and if the score is greater than 0.5, the mutation primarily stabilizes the protein; a DDG scores ≥ -0.5 and ≤ 0.05 represents a weak effect on protein stability [59]. iStable has two input types: structural and sequential. In the absence of structure, one can utilize the sequential input option for the stability prediction of the protein. After evaluating several machine learning algorithms, iStable adopted an SVM and had provided more accurate stability predictions compared to those of other tools [60]. Moreover, the HOPE (http://www.cmbi.ru.nl/hope) server was utilized, which provides insight into the structural and functional impact of a mutation. The protein sequence, wild type sequence, and substituted residue are submitted as inputs. HOPE server automatically accumulates and merges existing data from a series of web servers and databases and produces a mutation report complete with results, figures, and animations [61].

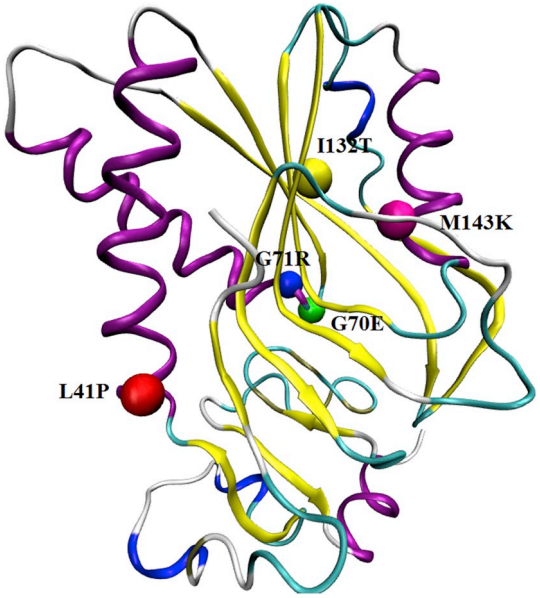

Fig. 1. Illustration of three-dimensional structures of the HPRT protein with mapped mutations.

2.3. Modeling of mutant proteins

Missense mutations predicted to be pathogenic as well as destabilizing were modeled onto the HPRT crystal structure (PDB ID: 1BZY) using Swiss PDB Viewer [62]. To remove any steric hindrance or clashes in the structure, we energy-minimized each structure using the same tool. Finally, the mutations were visualized using PyMOL software (Fig. 1).

2.4. Molecular dynamics simulation

The molecular dynamics simulation was carried out using Gromacs5.1.4 [63] for wild-type and mutant protein structures. The energy-minimized structure was obtained from the Swiss PDB viewer [62] and used as the starting structure for MD simulation. The wild-type and mutant protein structures were subjected to all-atom optimized potentials for the liquid simulation force field, which is an integral part of the Gromacs package. The gene box tool was used to set a cubic box at approximately 1 nm, and SPC water molecules were utilized for solvation. There are three necessary steps in molecular dynamics simulation: energy minimization, equilibrium, and MD production. The system was subjected to an initial energy minimization after the addition of ions using the steepest descent method, and this was continued until the system attained a minimum of 1000 kJ/mol. The equilibrium of the system further consisted of the isobaric-isothermal ensemble (NPT) and the canonical ensemble (NVT). Finally, a 50 ns molecular dynamics simulation was performed to analyze the various changes in the wild-type and mutant proteins.

3. Results and discussion

We used a series of in silico tools to predict the pathogenicity and stability of the publicly available nsSNPs of HPRT protein. SIFT, PolyPhen2, SNAP, and FATHMM were utilized to predict the pathogenicity and I-mutant3.0, and iStable were used to predict the stability as a consequence of missense mutations. Out of 88 nsSNPs, the SIFT tool predicted six nsSNPs found to be deleterious, with a score range of 0.02–0.03, and 67 nsSNPs were found to be highly deleterious, with a score equal to zero. To validate the results obtained from SIFT; we utilized PolyPhen-2, which uses sequence homology as well as structural information to predict the deleterious nature of nsSNPs. From 88

Table 1
Screening of deleterious missense SNPs associated to HPRT protein.

rs Id or Variant Id	Variants	SIFT	PolyPhen2.0	SNAP	FATHMM
VAR_006751	V8G	0	0.022	E	Damaging
VAR_006752	**G16D**	**0**	**0.993**	**E**	**Damaging**
rs137852480	**L41P**	**0**	**1**	**E**	**Damaging**
VAR_006757	**I42F**	**0**	**1**	**E**	**Damaging**
VAR_006758	**I42T**	**0**	**1**	**E**	**Damaging**
VAR_071611	**D44Y**	**0**	**1**	**E**	**Damaging**
VAR_006760	**R45K**	**0**	**1**	**E**	**Damaging**
VAR_006763	**A50P**	**0**	**1**	**E**	**Damaging**
VAR_006762	**A50V**	**0**	**1**	**E**	**Damaging**
VAR_006765	**R51P**	**0**	**0.953**	**E**	**Damaging**
VAR_006769	M54L	0	0	N	Damaging
Rs:137852495	M57T	0	0.009	N	Damaging
VAR_071613	A64P	0	0.971	E	Damaging
VAR_071614	**L65P**	**0**	**1**	**E**	**Damaging**
VAR_006773	**G70E**	**0**	**1**	**E**	**Damaging**
VAR_006774	**G71R**	**0**	**0.971**	**E**	**Damaging**
VAR_071615	**Y72C**	**0**	**0.267**	**E**	**Damaging**
rS137852481	**F74L**	**0**	**0.999**	**E**	**Damaging**
VAR_071616	**L78Q**	**0**	**1**	**E**	**Damaging**
rs137852483	**V130D**	**0**	**1**	**E**	**Damaging**
VAR_006781	**L131S**	**0**	**1**	**E**	**Damaging**
VAR_006783	**I132T**	**0**	**0.998**	**E**	**Damaging**
rs1378524896	**M143K**	**0**	**1**	**E**	**Damaging**
VAR_071619	**L147P**	**0**	**1**	**E**	**Damaging**
VAR_071620	**K159E**	**0**	**0.659**	**E**	**Damaging**
VAR_006788	**S162R**	**0**	**0.96**	**E**	**Damaging**
VAR_006790	**P176L**	**0**	**0.994**	**E**	**Damaging**
VAR_006791	**D177V**	**0**	**1**	**E**	**Damaging**
VAR_006792	**D177Y**	**0**	**1**	**E**	**Damaging**
VAR_006795	V188A	0	0.326	N	Damaging
rs267606863	**D194N**	**0**	**0.973**	**E**	**Damaging**
rs137852486	**F199V**	**0**	**0.985**	**E**	**Damaging**
VAR_006803	**D201Y**	**0**	**1**	**E**	**Damaging**
rs137852490	**H204D**	**0**	**0.747**	**E**	**Damaging**
VAR_006805	**H204R**	**0**	**1**	**E**	**Damaging**
VAR_006806	**C206Y**	**0**	**1**	**E**	**Damaging**
VAR_006750	G7D	0.25	0.023	E	Damaging
VAR_006753	G16S	0	0.792	N	Damaging
VAR_006754	D20V	0	0.417	E	Damaging
VAR_071610	**C23F**	**0**	**0.979**	**E**	**Damaging**
VAR_006755	**C23W**	**0**	**0.999**	**E**	**Damaging**
VAR_006761	**R48H**	**0**	**1**	**E**	**Damaging**
VAR_006764	**R51G**	**0**	**0.081**	**E**	**Damaging**
VAR_006767	V53A	0.03	0.218	E	Damaging
VAR_006768	V53M	0.02	0.89	N	Damaging
VAR_006771	G58R	0	0.029	N	Damaging
VAR_071612	H60R	0	0	N	Damaging
VAR_006776	L78V	0	0.991	N	Damaging
VAR_006777	D80V	0	0.732	N	Damaging
VAR_006778	**S104R**	**0**	**1**	**E**	**Damaging**
VAR_006779	**S110L**	**0**	**0.987**	**E**	**Damaging**
VAR_071618	T124P	0.03	0	N	Damaging
VAR_006782	**I132M**	**0**	**0.959**	**E**	**Damaging**
VAR_006784	**D135G**	**0**	**0.943**	**E**	**Damaging**
VAR_006787	A161S	0	0.998	N	Damaging
VAR_006789	T168I	0	0.258	N	Damaging
VAR_006796	I183T	0.03	0.902	N	Damaging
VAR_071622	**D185G**	**0**	**0.912**	**E**	**Damaging**
VAR_006795	V188A	0	0.326	N	Damaging
VAR_071623	**A192V**	**0**	**0.999**	**E**	**Damaging**
VAR_006797	D194E	0	0.52	E	Damaging
VAR_006799	**Y195C**	**0**	**0.999**	**E**	**Damaging**
VAR_006801	**D201G**	**0**	**0.951**	**E**	**Damaging**
VAR_006802	**D201N**	**0**	**0.863**	**E**	**Damaging**
424675	G16V	0.17	0.997	E	Damaging
422413	R34G	0.35	0	N	Damaging
38940	**R48H**	**0**	**1**	**E**	**Damaging**
264805	**L49F**	**0.04**	**1**	**E**	**Damaging**
25098	R51G	0.36	0.081	E	Damaging
25118	**D52G**	**0.02**	**0.869**	**E**	**Damaging**
265138	**L68F**	**0**	**0.978**	**E**	**Damaging**
430651	Q109K	1.0	0.001	N	Damaging
446501	L122F	0.12	0.633	N	Damaging
25119	**G140D**	**0**	**1**	**E**	**Damaging**

(*continued on next page*)

Table 1 (*continued*)

rs Id or Variant Id	Variants	SIFT	PolyPhen2.0	SNAP	FATHMM
580774	R151S	0.59	0	N	Damaging
266922	S162I	0.16	0.99	E	Damaging
98513	**K166E**	**0**	**0.79**	**E**	**Damaging**
551342	V179A	0.24	0.011	E	Damaging
422414	**V189A**	**0**	**1**	**E**	**Damaging**
268410	**D194H**	**0**	**0.898**	**E**	**Damaging**
25086	H204D	0.08	0.747	E	Damaging
87575	D18E	0.60	0	N	Damaging
87636	N26T	0.60	0	N	Damaging
86562	**R48C**	**0**	**0.849**	**E**	**Damaging**
64779	Q109L	0.25	0.195	N	Damaging
87757	P169A	0.94	0.006	N	Damaging
87756	R170Q	0.50	0.647	E	Damaging
87527	K217R	0.35	0	N	Damaging

Bold highlighted mutations are found to be deleterious from all the tools.

nsSNPs, PolyPhen2 predicted 56 nsSNPs as likely damaging, with a range of 0.85–1, 17 nsSNPs as possibly damaging, within a score range of 0.02–0.85, and the remaining 15 nsSNPs as benign, with a score range of 0–0.02. SNAP predicted that 69 nsSNPs would affect the structure of the protein and that 19 nsSNPs were neutral. The hidden Markov model-based FATHMM tool predicted all nsSNPs to be damaging. After combining the results of all the tools (SIFT, PolyPhen2, SNAP, and FATHMM), a total of 53 nsSNPs from 88 nsSNPs were predicted to be deleterious (Table 1).

The filtered dataset of 53 nsSNPs was further subjected to stability analysis using I-mutant3.0 and iStable. Out of 53 nsSNPs, 47 were predicted to decrease protein stability and 5 nsSNPs to increase protein stability by I-mutant3.0. iStable predicted that 42 nsSNPs would decrease protein stability and that the remaining 11 nsSNPs would increase protein stability. In combination (I-mutant3.0 and iStable), 40 nsSNPs were found to reduce protein stability (Table 2). Based on the clinical pathogenicity and association with LND, we selected five nsSNPs (L41P, G70E, G71R, I132T, and M143K) that were found to be deleterious and destabilizing by all the by all in silico prediction tools for further analysis.

Moreover, we utilized the HOPE server to calculate the physicochemical properties of the selected mutations. L41P and I132T introduce a smaller residue that causes space in the core of the protein, whereas G70E, G71R, and M143K introduce a larger residue that probably will not fit in the protein structure. The wild-type residues G70E, G71R, I132T, and M143K, are more hydrophobic. Hence, a mutation at these positions will disrupt significant hydrophobic interactions. In both the PDB file and in the PISA assembly, the L41P residue was found to be involved in a multimer contacts. The PISA database contains protein assemblies that are highly likely to be biologically relevant. This is a strong indication that the residue is indeed in contact with other proteins. The new residue might be too small to make multimer contacts. The wild-type residue G70E in the 3D-structure has interactions with a ligand annotated as POP. The difference in properties between the wild-type and mutant protein can easily cause a loss of interactions with a ligand. Because ligand binding is often essential for protein function, this function might be disturbed by the G70E mutation. The G70E mutant residue introduces a negative charge in a buried residue, which can lead to protein folding problems.

G71R substitutes a positively charged residue for a neutral residue. The wild-type residue glycine is the most flexible of all the residues, and the flexibility might be necessary for the protein's function. Mutation of this glycine could abolish this function. The torsion angles for this residue are unusual. Only glycine is flexible enough to make these torsion angles; mutation into another residue may force the local backbone into an incorrect conformation, thus disturbing the local structure. M143K also introduces a positive charge in a buried residue, which can lead to protein folding problems. The mutated residue is not in direct contact with a ligand. However, the mutation could affect local stability, which in turn could affect ligand contacts made by one of the neighboring residues.

Table 2
Stability prediction analysis of missense SNPs associated to HPRT protein.

rs Id or Variant Id	Variants	I-Mutant3.0	Istable
VAR_006752	**G16D**	**Decrease**	**Decrease**
rs137852480	**L41P**	**Decrease**	**Decrease**
VAR_006757	**I42F**	**Decrease**	**Decrease**
VAR_006758	**I42T**	**Decrease**	**Decrease**
VAR_071611	**D44Y**	**Decrease**	**Decrease**
VAR_006760	R45K	Decrease	Increase
VAR_006763	**A50P**	**Decrease**	**Decrease**
VAR_006762	A50V	Decrease	Increase
VAR_006765	**R51P**	**Decrease**	**Decrease**
VAR_071614	**L65P**	**Decrease**	**Decrease**
VAR_006773	**G70E**	**Decrease**	**Decrease**
VAR_006774	**G71R**	**Decrease**	**Decrease**
VAR_071615	**Y72C**	**Decrease**	**Decrease**
rS137852481	F74L	Increase	Decrease
VAR_071616	**L78Q**	**Decrease**	**Decrease**
rs137852483	**V130D**	**Decrease**	**Decrease**
VAR_006781	L131S	Decrease	Increase
VAR_006783	**I132T**	**Decrease**	**Decrease**
rs1378524896	**M143K**	**Decrease**	**Decrease**
VAR_071619	**L147P**	**Decrease**	**Decrease**
VAR_071620	K159E	Increase	Decrease
VAR_006788	S162R	Decrease	Increase
VAR_006790	P176L	Decrease	Increase
VAR_006791	**D177V**	**Decrease**	**Decrease**
VAR_006792	**D177Y**	**Decrease**	**Decrease**
rs267606863	**D194N**	**Decrease**	**Decrease**
rs137852486	**F199V**	**Decrease**	**Decrease**
VAR_006803	D201Y	Decrease	Increase
rs137852490	H204D	Decrease	Increase
VAR_006805	**H204R**	**Decrease**	**Decrease**
VAR_006806	**C206Y**	**Decrease**	**Decrease**
VAR_071610	**C23F**	**Decrease**	**Decrease**
VAR_006755	**C23W**	**Decrease**	**Decrease**
VAR_006761	**R48H**	**Decrease**	**Decrease**
VAR_006764	R51G	Decrease	Increase
VAR_006778	S104R	Increase	Decrease
VAR_006779	**S110L**	**Decrease**	**Decrease**
VAR_006782	I132M	Increase	Decrease
VAR_006784	**D135G**	**Decrease**	**Decrease**
VAR_071622	D185G	Decrease	Increase
VAR_071623	**A192V**	**Decrease**	**Decrease**
VAR_006799	**Y195C**	**Decrease**	**Decrease**
VAR_006801	**D201G**	**Decrease**	**Decrease**
VAR_006802	**D201N**	**Decrease**	**Decrease**
38940	**R48H**	**Decrease**	**Decrease**
264805	**L49F**	**Decrease**	**Decrease**
25118	**D52G**	**Decrease**	**Decrease**
265138	**L68F**	**Decrease**	**Decrease**
25119	**G140D**	**Decrease**	**Decrease**
98513	K166E	Increase	Increase
422414	**V189A**	**Decrease**	**Decrease**
268410	**D194H**	**Decrease**	**Decrease**
86562	**R48C**	**Decrease**	**Decrease**

IDs highlighted in bold are found to be destabilizing from all the tools.

3.1. Molecular dynamic simulation analysis

In recent years, MD simulation data have proven to be a powerful approach in understanding macromolecular structure-function relationships [64–66]. Here, we performed a 50 ns MD simulation for all the mutant proteins and compared the results with those for the wild-type protein. Trajectories are sequential snapshots of the simulated molecular system, which represents atomic coordinates at specific time periods. To investigate the molecular structural conformation and functional changes of wild-type and mutant HPRT proteins, and its association with LND the resultant trajectories were subjected to

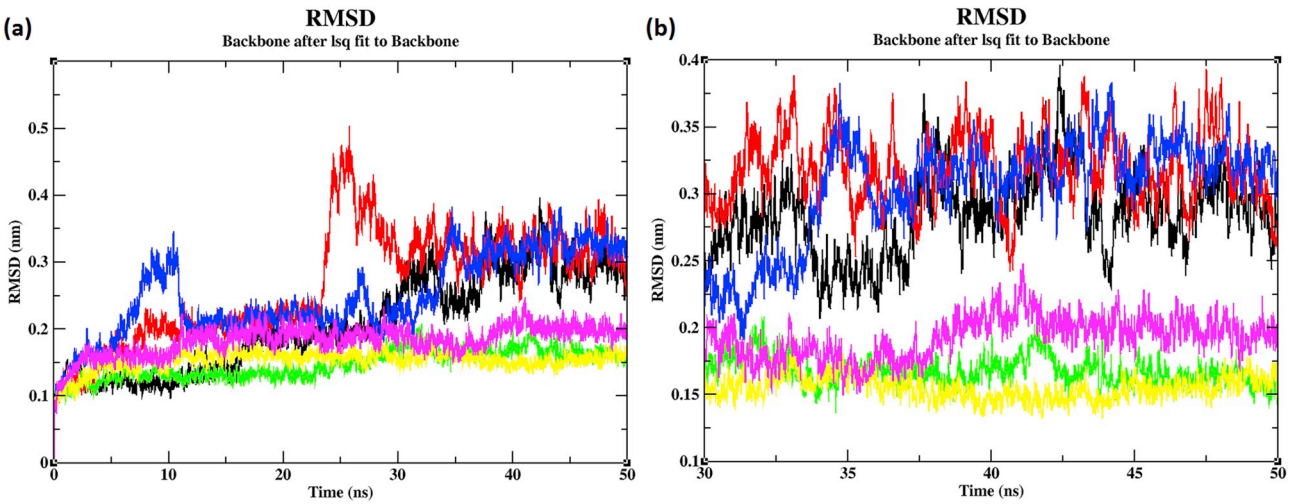

Fig. 2. The root mean square deviation (RMSD) plot of wild-type and mutant HPRT over 50 ns of simulation [Color scheme is as follows: wild-type (black color), L41P (red color), G70E (green color), G71R (blue color), I132T (yellow color), and M143K (Magenta color)]. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

various analyses. This involves stability and functional movements prediction analysis using root mean square deviation (RMSD) and root mean square fluctuation (RMSF), compactness analysis using radius of gyration (Rg), hydrophilic and hydrophobic analysis using solvent accessible surface area (SASA), and collective motion analysis of the protein atoms using covariance matrix. The principal component analysis allows us to understand the highest stable state of the protein using the free energy landscape and significant structural changes in the secondary structure elements using secondary structure element analysis.

3.2. Root mean square deviation

RMSD was calculated to predict the stability of the protein. The RMSD of the C alpha atoms was estimated using the 50 ns MDS trajectory of the wild-type and mutant proteins (L41P, G70E, G71R, I132T, and M143K). The RMSD plot demonstrated that both the wild-type and mutant proteins attained a stable conformation after 30 ns (Fig. 2a). Thus, we selected the trajectories between 30 ns and 50 ns for further analysis (Fig. 2b). When compared with the wild-type protein, mutant proteins containing G70E, I132T, and M143K showed lower RMSD values, whereas mutants proteins containing L41P and G71R showed slightly higher RMSD values (Fig. 2b). A higher RMSD value signifies a decrease in protein stability [42]; conversely, a lower RMSD value characterizes an increase in protein stability [43]. In this context, mutants G70E, I132T, and M143K stabilize the protein, and L41P and G71R destabilize the protein structure.

3.3. Root mean square fluctuation

To validate the RMSD outcomes, we calculated the RMSF, which provides details of each flexible residue within the protein structure. RMSF values of the wild-type and mutant proteins were compiled to determine whether mutations affect the dynamic behavior of residues. The backbone RMSF of the wild-type residue and each mutant residue of HPRT was calculated to illustrate the flexibility of the structure. Higher RMSF values imply more flexible movements, while lower RMSF values correspond to limited movements during MD simulation concerning the average position. Higher flexibility leads to protein destabilization [41,42]. The RMSF plot (Fig. 3a) indicates that the mutants L41P, G70E, G71R, I132T, and M143K have lower RMSF values than that of the wild-type, whereas mutants L41P, G70E, and G71R showed higher fluctuations. In mutant L41P, residues 40 to 55, 78 to 90 and 168 to 177 showed higher fluctuations (Fig. 3a). However, the mutations G70E and G71R exhibited higher flexibility between residues 80 to 90 and 160 to 180 and residues 105 to 120 and 168 to 175, respectively (Fig. 3a). These significant fluctuation patterns and their associated functional changes were further studied by a correlation matrix and PCA analysis.

3.4. Compactness analysis

Rg is a crucial parameter for assessing the dynamic adaptability of a protein in a solution and provides details of the compactness of a protein by measuring the mass-weighted root mean square distance of the group of atoms from their common center of mass. A stably folded protein maintains a stable Rg value; however, during protein unfolding, Rg changes over time. A lower Rg value suggests better compactness of the protein [41,43,67]. A clear difference in the Rg pattern was observed between the wild-type and mutant proteins (Fig. 3b). In comparison to the wild-type protein, the mutants L41P and G71R displayed slightly lower Rg values, while mutants G70E, I132T, and M143K demonstrated much lower Rg values (Fig. 3b). The outcome of the Rg analysis suggests that among the mutants, L41P and G71R increase the compactness subtly, whereas G70E, I132T, and M143K increase the compactness at a higher level.

3.5. Solvent accessible surface area

SASA describes solvent accessible residues by analyzing the hydrophobic and hydrophilic residues of a protein. A change in SASA manifests a change in protein structure and consequently protein function [39,42,43]. An alteration in SASA between the wild-type and mutant proteins was observed. In comparison to the wild-type, mutant G70E (Fig. 4a) showed slightly higher SASA value, while mutants L41P (Fig. 4a), I132T, and M143K (Fig. 4b) showed lower SASA values. Mutant G71R and the wild-type protein showed relatively similar SASA values (Fig. 4b). The SASA results indicate that the structural rearrangements in the mutant G70E resulted in increased solvent accessibility, whereas conversely, structural rearrangements in the mutants L41P, I132T, and M143K decreased solvent accessibility. SASA clearly dictates the structural rearrangements in the L41P, G70E, I132T, and M143K mutants, thus suggesting that these mutations significantly change the function of the HPRT protein.

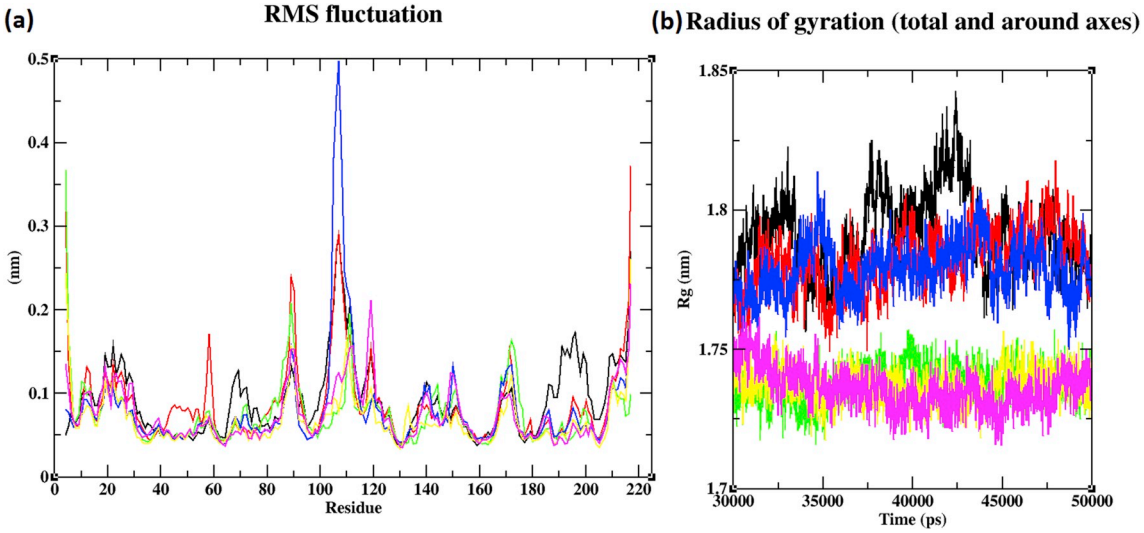

Fig. 3. The root mean square fluctuation (RMSF) (a) and Rg (b) plot of wild-type and mutant HPRT over 50 ns of simulation [Color scheme is as follows: wild-type (black color), L41P (red color), G70E (green color), G71R (blue color), I132T (yellow color), and M143K (Magenta color)]. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

3.6. Covariance analysis

A covariance matrix calculates collective motion or essential dynamics by examining atoms in cooperation instead of focusing on more local fluctuations. Collective motions reveal the most significant biological functions, such as conformational change, protein-protein interactions, protein-ligand interactions, dynamic stability, globular and membrane proteins, pore opening and closing, ion gating, signal transduction, or substrate translocation via membrane protein, which are facilitated by the cooperative movements of symmetrically organized subunits. These findings suggest that the understanding of correlated motion is essential to gain insights into structure-function relationships [68]. Covariance matrix color spectrum labeling using cyan and light green represents strong and very weak or no correlation of atom pairs, respectively. In the covariance matrix of the wild-type protein, few strong correlations were spotted that were encircled in red; each circle was assigned a number for a better comparison with the mutants. The correlations observed in the wild-type (Fig. 5a) were diminished in strength or completely absent in all the mutants (L41P (Fig. 5b), G70E (Fig. 5c), G71R (Fig. 5d), I132T (Fig. 5e), and M143K (Fig. 5f); aside from these, the mutants L41P (Fig. 5b) and G71R (Fig. 5d) retained the 4th correlation. These variations of correlated movements of atom pairs suggested a change in the structural conformation of the mutants that could potentially alter the function of the mutant HPRT protein.

3.7. Principal component analysis

PCA utilizes covariance outputs to compute aggregate global motions in an essential subspace. PCA eases the computation of complex,

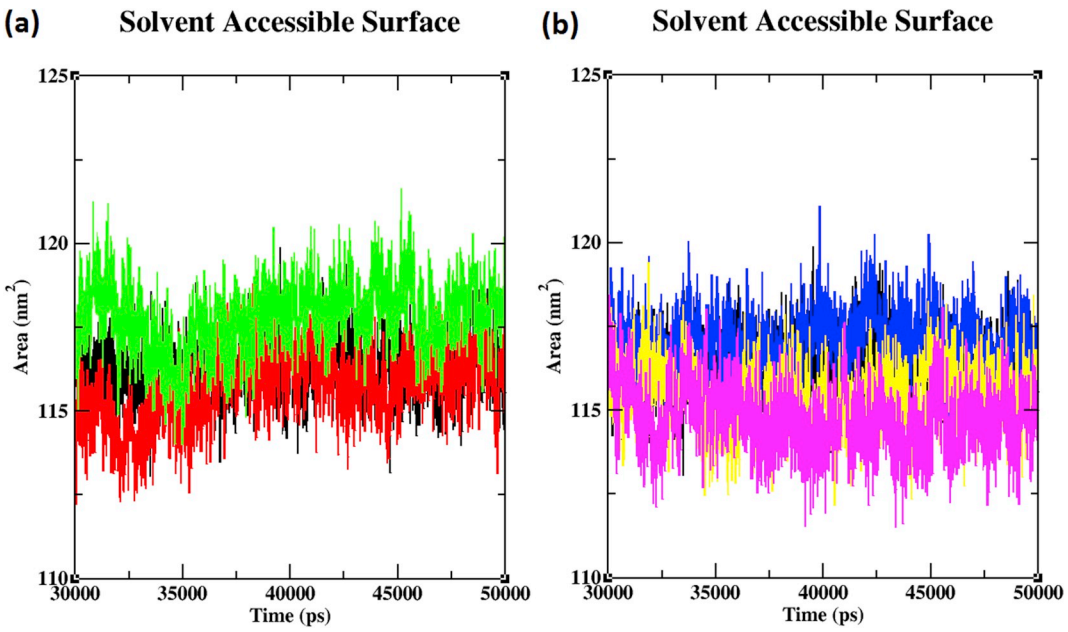

Fig. 4. The solvent accessible surface area (SASA) plot of wild-type and mutant HPRT over 50 ns of simulation [Color scheme is as follows: wild-type (black color), L41P (red color), G70E (green color), G71R (blue color), I132T (yellow color), and M143K (Magenta color)]. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

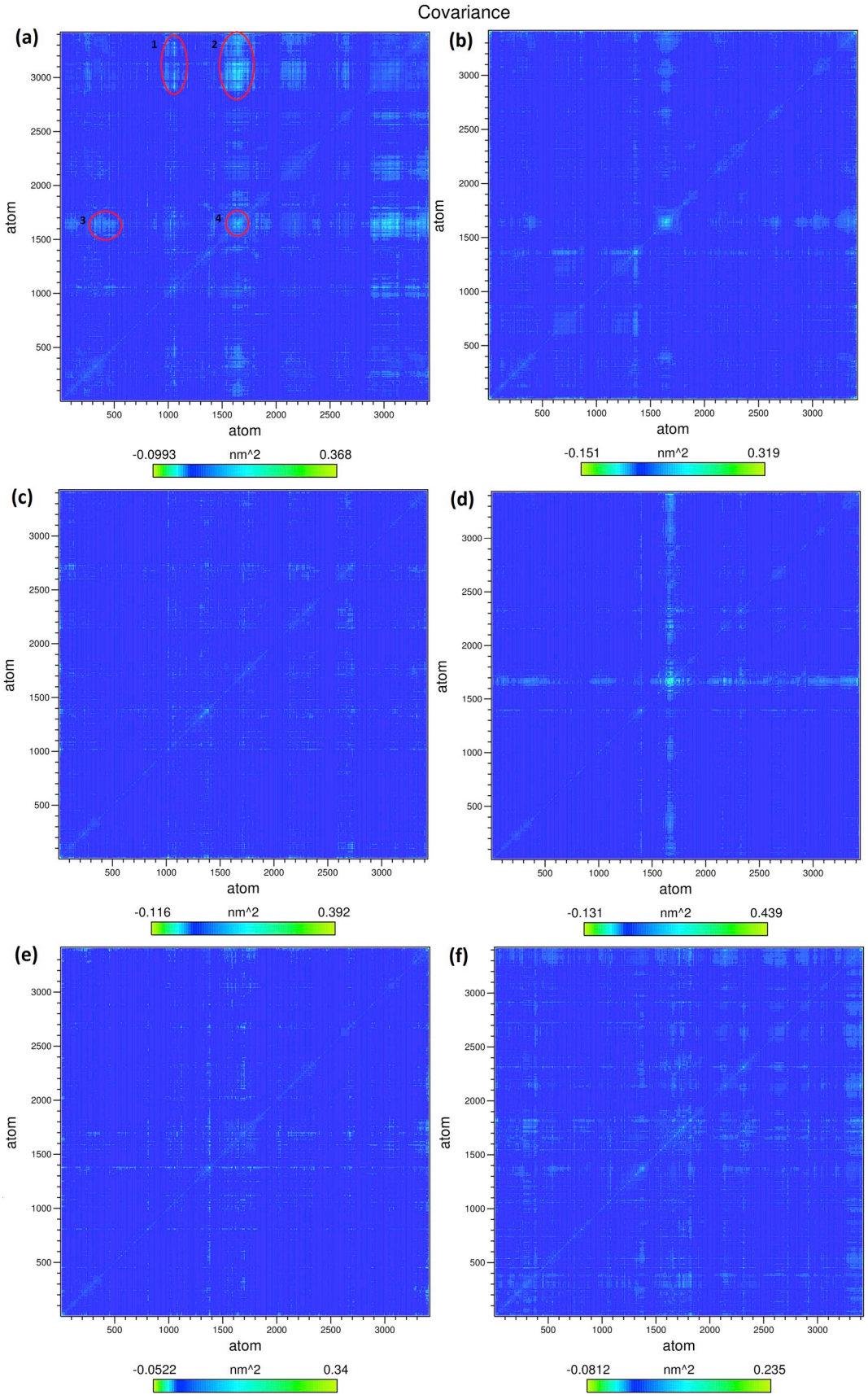

Fig. 5. The covariance analysis of wild-type and mutant HPRT over 50 ns of simulation (a) wild-type, (b) L41P, (c) G70E, (d) G71R, (e) I132T, and (f) M143K.

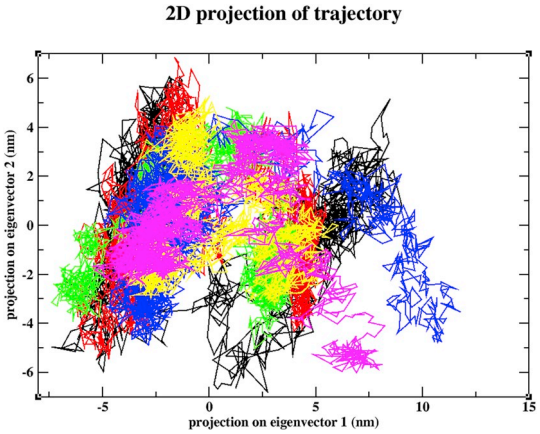

Fig. 6. The principal component analysis (PCA) of wild-type and mutant HPRT over 50 ns of simulation [Color scheme is as follows: wild-type (black color), L41P (red color), G70E (green color), G71R (blue color), I132T (yellow color), and M143K (Magenta color)]. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

significant fluctuations of atom pairs that are related to crucial biological functions [69]. The principal component motions of the wild-type and mutant HPRT protein were calculated using PCA. The first two principal components (PC1 & PC2) were selected to analyze collective motions. Principal component 1 (PC1) records the direction of most variations, and principal component 2 (PC2) records the direction of the second most variations. The fluctuations of the wild-type protein and mutant G71R covered a comparatively similar subspace, while other mutants (L41P, G70E, I132T, and M143K) covered the lower subspace along with a projection of eigenvector 1 and projection of eigenvector 2 (Fig. 6). The results of the PCA signify that mutant G71R has less of an impact, while mutants L41P, G70E, I132T, and M143K have higher impacts regarding changing the structural conformation and collective motion behavior of the HPRT protein.

3.8. Secondary structure analysis

Secondary structure exploration provides the details of significant structural changes in the secondary structure elements of a protein [70]. A secondary structure plot (Fig. 7) revealed that the wild-type protein (Fig. 7a) and mutant L41P (Fig. 7b) have almost the same secondary structure element conformation. However, the mutants G70E

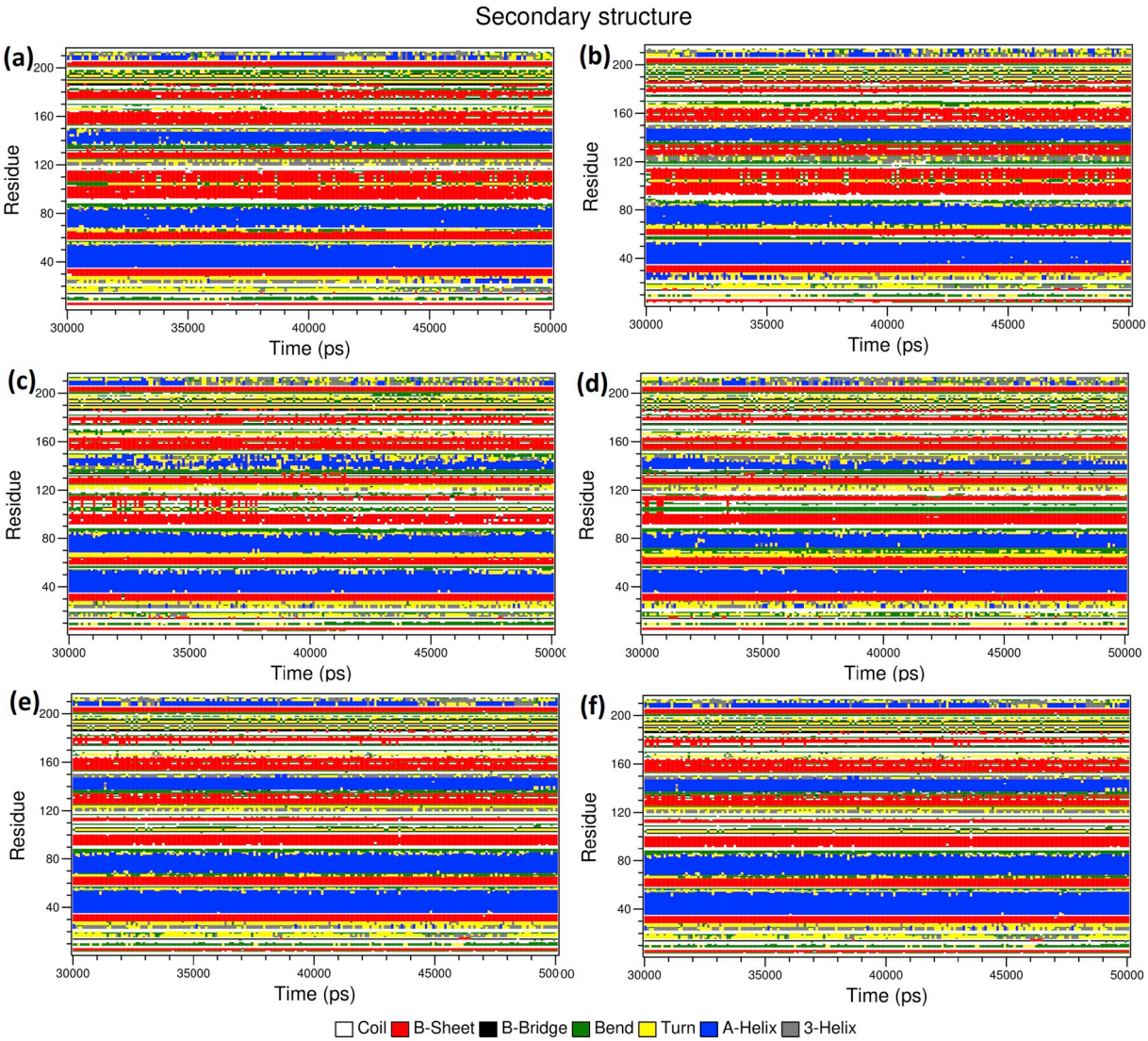

Fig. 7. The secondary structure analysis of wild-type and mutant HPRT over 50 ns of simulation (a) wild-type, (b) L41P, (c) G70E, (d) G71R, (e) I132T, and (f) M143K.

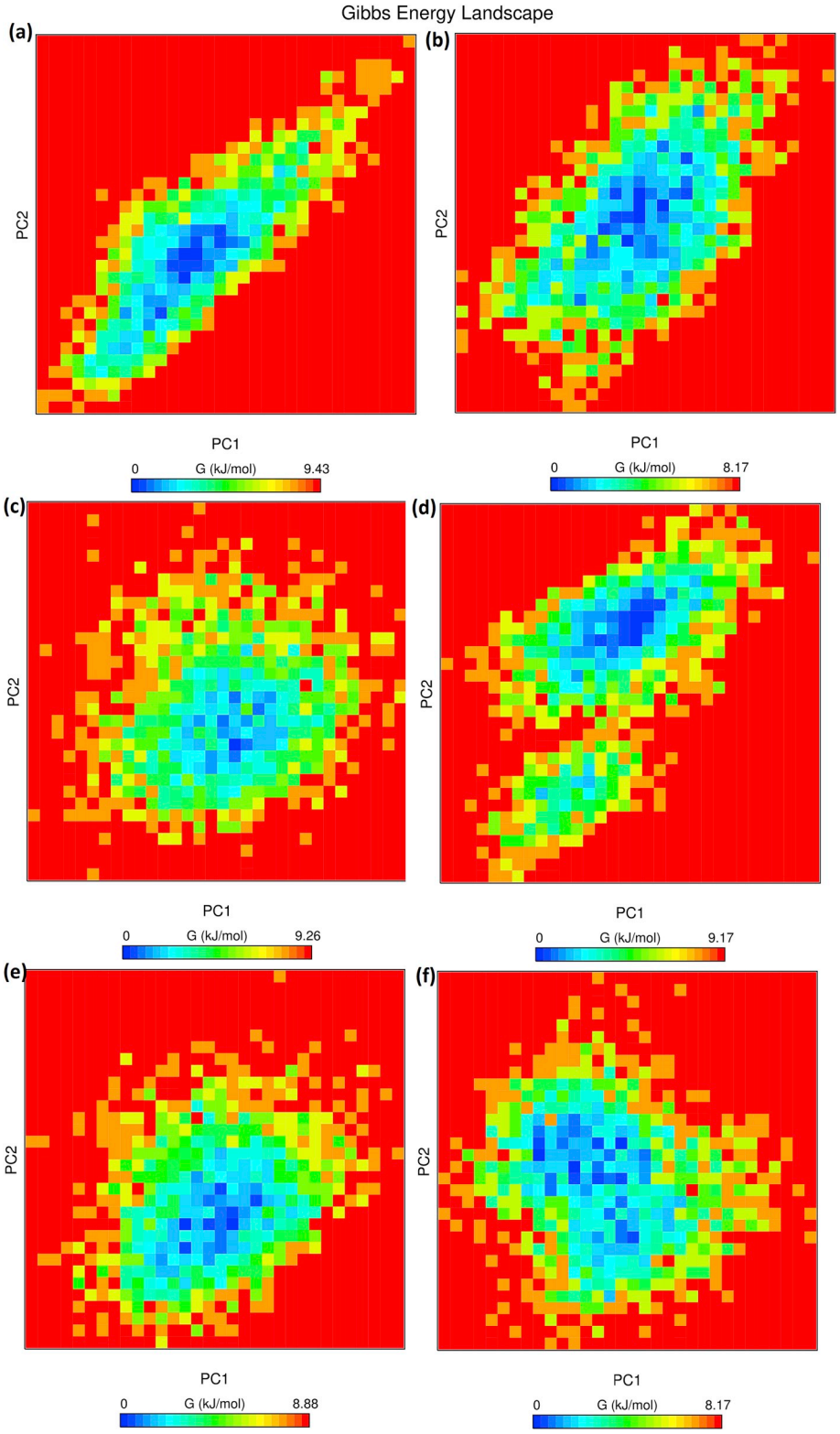

Fig. 8. The free energy landscape (FEL) analysis of wild-type and mutant HPRT over 50 ns of simulation (a) wild-type, (b) L41P, (c) G70E, (d) G71R, (e) I132T, and (f) M143K.

(Fig. 7c), G71R (Fig. 7d), I132T (Fig. 7e), and M143K (Fig. 7f) showed a transition from beta-sheet to bend, turn and coil at residues 105 to 115; the other elements were marginally similar to the wild-type protein. Moreover, the mutant G71R displayed a transition from helix to bend and turned in the vicinity of the mutant position (at residues 69–72), which is in concordance with the Chou-Fasman secondary structure analysis prediction. This analysis of the wild-type (Fig. 7a) and mutant proteins indicates that mutant L41P (Fig. 7b) retains secondary structure elements intact; conversely, mutants G70E (Fig. 7c), G71R (Fig. 7d), I132T (Fig. 7e), and M143K (Fig. 7f) show damage to the beta-sheet conformation at residues 105 to 115, which might be an essential component for HPRT protein function.

3.9. Free energy landscape (FEL)

FEL analysis (Fig. 8) provides insight into folding patterns in terms of Gibbs free energy change, a pivotal aspect for determining protein function. To comprehend the folding pattern from a FEL plot, the first two principal components of atomic conformation rerecorded along with their corresponding Gibbs free energy [71,72]. The FEL color spectrum labeling ranges from blue to red; here, blue corresponds to the subspace of least energy with a global minima conformation that represents the highest stable state of the protein, and the red color corresponds to a subspace of high energy, i.e., a highly unstable stable state of the protein. Between the blue and red, the cyan-green and yellow color spectrum signify an intermediary energy state, which defines the energy barrier between the highly stable and highly unstable conformation of the protein. The wild-type (Fig. 8a) and mutants L41P (Fig. 8b), G70E (Fig. 8c), I132T (Fig. 8e), and M143K (Fig. 8f) displayed one global energy minima along with the first two principal components. However, mutant G71R (Fig. 8d) presented one global energy minima as well as one local energy minima. All mutants exhibited a high energy barrier between highly stable and highly unstable conformations in comparison to the wild-type protein. The FEL results indicate that because of the high energy barrier, all the mutants have difficulty of achieving the global minima conformation, in contrast to the wild-type HPRT protein.

4. Conclusion

In the present study, we used a series of in silico tools in combination with MDS to filter the most deleterious missense mutations that affect the protein structure and function of the HPRT protein. We investigated numerous molecular properties, including trajectory stabilization, atomic flexibility, compactness, solvent accessibility, global aggregate motions, folding pattern, and transitions in secondary structure elements, of the wild-type protein and all the mutants. We conclude that five mutants (L41P, G70E, G71R, I132T, and M143K) alter the structural conformation and dynamic stability of the HPRT protein and are considered to be the most deleterious based on our combinatorial computational analyses. These findings corroborate the level pathogenicity of the HPRT protein mutants and their associations with the phenotype of patients with LND. The computational pipeline that we used is expected to serve as a potential way to prioritize potential targets that could be used as therapeutics for LND.

Acknowledgment

We want to thank VIT management for providing the facility, seed money, and platform to carry out the research. We also would like to acknowledge Prof. Chiranjib Chakraborty for his technical assistance. The authors declare that no conflicts of interest exist.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compbiomed.2019.02.014.

References

[1] W.L. Nyhan, W.J. Oliver, M. Lesch, A familial disorder of uric acid metabolism and central nervous system function. II, J. Pediatr. 67 (1965) 257–263, https://doi.org/10.1016/S0022-3476(65)80249-9.
[2] J.E. Seegmiller, F.M. Rosenbloom, W.N. Kelley, Enzyme defect associated with a sex-linked human neurological disorder and excessive purine synthesis, Science 155 (80) (1967) 1682–1684, https://doi.org/10.1126/science.155.3770.1682.
[3] K.V. Nguyen, R.K. Naviaux, W.L. Nyhan, Novel mutation in the human HPRT1 gene and the Lesch-Nyhan disease, Nucleos Nucleot. Nucleic Acids (2017), https://doi.org/10.1080/15257770.2017.1395037.
[4] S. Guillemette, A. Branagan, M. Peng, A. Dhruva, O.D. Scharer, S.B. Cantor, FANCJ localization by mismatch repair is vital to maintain genomic integrity after UV irradiation, Cancer Res. (2014), https://doi.org/10.1158/0008-5472.CAN-13-2474.
[5] M. Faridounnia, H. Wienk, L. Kovačič, G.E. Folkers, N.G.J. Jaspers, R. Kaptein, J.H.J. Hoeijmakers, R. Boelens, The cerebro-oculo-facio-skeletal syndrome point mutation F231L in the ERCC1 DNA repair protein causes dissociation of the ERCC1-XPF complex, J. Biol. Chem. (2015), https://doi.org/10.1074/jbc.M114.635169.
[6] C.W. Op het Veld, S. van Hees-Stuivenberg, A.A. van Zeeland, J.G. Jansen, Effect of nucleotide excision repair on hprt gene mutations in rodent cells exposed to DNA ethylating agents, Mutagenesis 12 (6) (1997) 417–424.
[7] R.K. Chiu, J. Brun, C. Ramaekers, J. Theys, L. Weng, P. Lambin, D.A. Gray, B.G. Wouters, Lysine 63-polyubiquitination guards against translesion synthesis-induced mutations, PLoS Genet. (2006), https://doi.org/10.1371/journal.pgen.0020116.
[8] B.L. Davidson, M. Pashmforoush, W.N. Kelley, T.D. Palella, Human hypoxanthine-guanine phosphoribosyltransferase deficiency. The molecular defect in a patient with gout (HPRT(ASHVILLE)), J. Biol. Chem. 264 (1989) 520–525.
[9] W.N. Kelley, F.M. Rosenbloom, J.F. Henderson, J.E. Seegmiller, A specific enzyme defect in gout associated with overproduction of uric acid, Proc. Natl. Acad. Sci. U. S. A 57 (1967) 1735–1739 http://www.ncbi.nlm.nih.gov/pubmed/4291947 , Accessed date: 24 October 2018.
[10] P.I. Patel, P.E. Framson, C.T. Caskey, A.C. Chinault, Fine structure of the human hypoxanthine phosphoribosyltransferase gene, Mol. Cell Biol. 6 (1986) 393–403, https://doi.org/10.1177/201010581202105101.
[11] R. Fu, L. Ceballos-Picot, J. Rosa, L.E. Larovere, Y. Yamada, K.V. Nguyen, M. Hegde, J.E. Visser, D.J. Schretlen, W.L. Nyhan, J.G. Puig, P.J. O'Neill, H.A. Jinnah, Genotype-phenotype correlations in neurogenetics: lesch-Nyhan disease as a model disorder, Brain 137 (2014) 1282–1303, https://doi.org/10.1093/brain/awt202.
[12] H.A. Jinnah, L. De Gregorio, J.C. Harris, W.L. Nyhan, J.P. O'Neill, The spectrum of inherited mutations causing HPRT deficiency: 75 new cases and a review of 196 previously reported cases, Mutat. Res. Rev. Mutat. Res. 463 (2000) 309–326, https://doi.org/10.1016/S1383-5742(00)00052-1.
[13] H.A. Jinnah, J.C. Harris, W.L. Nyhan, J.P. O'Neill, The spectrum of mutations causing HPRT deficiency: an update, Nucleos Nucleot. Nucleic Acids (2004) 1153–1160, https://doi.org/10.1081/NCN-200027400.
[14] H.A. Jinnah, I. Ceballos-Picot, R.J. Torres, J.E. Visser, D.J. Schretlen, A. Verdu, L.E. Laróvere, C.J. Chen, A. Cossu, C.H. Wu, R. Sampat, S.J. Chang, R.D. De Kremer, W. Nyhan, J.C. Harris, S.G. Reich, J.G. Puig, Attenuated variants of Lesch-Nyhan disease, Brain 133 (2010) 671–689, https://doi.org/10.1093/brain/awq013.
[15] K.V. Nguyen, R.K. Naviaux, W.L. Nyhan, Human HPRT1 gene and the Lesch–Nyhan disease: substitution of alanine for glycine and inversely in the HGprt enzyme protein, Nucleos Nucleot. Nucleic Acids 36 (2017) 151–157, https://doi.org/10.1080/15257770.2016.1231319.
[16] R.J. Torres, J.G. Puig, GLUT 9 influences uric acid concentration in patients with Lesch-Nyhan disease, Int. J. Rheum. Dis. (2018), https://doi.org/10.1111/1756-185X.13323.
[17] B.L. Davidson, S.A. Tarle, T.D. Palella, W.N. Kelley, Molecular basis of hypoxanthine-guanine phosphoribosyltransferase deficiency in ten subjects determined by direct sequencing of amplified transcripts, J. Clin. Invest. 84 (1989) 342–346, https://doi.org/10.1172/JCI114160.
[18] Y. Yamada, K. Yamada, S. Sonta, N. Wakamatsu, N. Ogasawara, Mutations in the hypoxanthine guanine phosphoribosyltransferase gene (HPRT1) in Asian HPRT deficient families, Nucleos Nucleot. Nucleic Acids (2004) 1169–1172, https://doi.org/10.1081/NCN-200027439.
[19] S.A. Tarlé, B.L. Davidson, V.C. Wu, F.J. Zidar, J.E. Seegmiller, W.N. Kelley, T.D. Palella, Determination of the mutations responsible for the lesch-nyhan syndrome in 17 subjects, Genomics 10 (1991) 499–501, https://doi.org/10.1016/0888-7543(91)90341-B.
[20] Y. Yamada, N. Nomura, K. Yamada, N. Wakamatsu, Molecular analysis of HPRT deficiencies: an update of the spectrum of Asian mutations with novel mutations, Mol. Genet. Metabol. 90 (2007) 70–76, https://doi.org/10.1016/j.ymgme.2006.08.013.
[21] S. Fujimori, B.L. Davidson, W.N. Kelley, T.D. Palella, Identification of a single nucleotide change in the hypoxanthine-guanine phosphoribosyltransferase gene (HPRT(Yale)) responsible for Lesch-Nyhan syndrome, J. Clin. Invest. 83 (1989) 11–13, https://doi.org/10.1172/JCI113846.
[22] S. Fujimori, T. Yamanouchi, I. Akaoka, HPRT deficiency, Nihon Rinsho. 49 (1991) 1036–1042 http://www.ncbi.nlm.nih.gov/pubmed/2067091 , Accessed date: 24 September 2018.
[23] E. Zoref-Shani, Y. Bromberg, J. Hirsch, S. Feinstein, Y. Frishberg, O. Sperling, A novel point mutation (I137T) in the conserved 5-phosphoribosyl-1-pyrophosphate binding motif of hypoxanthine-guanine phosphoribosyltransferase (HPRTJerusalem) in a variant of Lesch-Nyhan syndrome, Mol. Genet. Metabol. 78 (2003) 158–161, https://doi.org/10.1016/S1096-7192(03)00002-7.
[24] R.A. Gibbs, P.N. Nguyen, A. Edwards, A.B. Civitello, C.T. Caskey, Multiplex DNA deletion detection and exon sequencing of the hypoxanthine phosphoribosyltransferase gene in Lesch-Nyhan families, Genomics 7 (1990) 235–244, https://doi.org/10.1016/0888-7543(90)90545-6.
[25] J.M. Wilson, J.T. Stout, T.D. Palella, B.L. Davidson, W.N. Kelley, C.T. Caskey, A molecular survey of hypoxanthine- guanine phosphoribosyl-transferase deficiency in man, J. Clin. Invest 77 (1986) 188–195, https://doi.org/10.1172/JCI112275.
[26] J.T.L. Mah, K.S. Chia, A gentle introduction to SNP analysis: resources and tools, J. Bioinf. Comput. Biol. 5 (2007) 1123–1138, https://doi.org/10.1142/S0219720007003090.
[27] A. Riva, I.S. Kohane, SNPper: retrieval and analysis of human SNPs, Bioinformatics 18 (2002) 1681–1685, https://doi.org/10.1093/bioinformatics/18.12.1681.
[28] Q. Ma, A.Y.H. Lu, Pharmacogenetics, pharmacogenomics, and individualized medicine, Pharmacol. Rev. 63 (2011) 437–459, https://doi.org/10.1124/pr.110.

003533.
[29] N.Z. Gerek, L. Liu, K. Gerold, P. Biparva, E.D. Thomas, S. Kumar, Evolutionary Diagnosis of non-synonymous variants involved in differential drug response, BMC Med. Genomics 8 (2015) S6, https://doi.org/10.1186/1755-8794-8-S1-S6.
[30] M. Krawczak, E. V Ball, I. Fenton, P.D. Stenson, S. Abeysinghe, N. Thomas, D.N. Cooper, Human gene mutation database-a biomedical information and research resource, Hum. Mutat. 15 (2000) 45–51, https://doi.org/10.1002/(SICI)1098-1004(200001)15:1<45::AID-HUMU10>3.0.CO;2-T.
[31] Z. Wang, J. Moult, SNPs, protein structure, and disease, Hum. Mutat. 17 (2001) 263–270, https://doi.org/10.1002/humu.22.
[32] S. Sunyaev, V. Ramensky, P. Bork, Towards a structural basis of human non-synonymous single nucleotide polymorphisms, Trends Genet. 16 (2000) 198–200, https://doi.org/10.1016/S0168-9525(00)01988-0.
[33] P. Sneha, T.U. Zenith, U.S. Abu Habib, J. Evangeline, D. Thirumal Kumar, C. George Priya Doss, R. Siva, H. Zayed, Impact of missense mutations in survival motor neuron protein (SMN1) leading to Spinal Muscular Atrophy (SMA): a computational approach, Metab. Brain Dis. (2018), https://doi.org/10.1007/s11011-018-0285-4.
[34] D. Thirumal Kumar, L. Jerushah Emerald, C. George Priya Doss, P. Sneha, R. Siva, W. Charles Emmanuel Jebaraj, H. Zayed, Computational approach to unravel the impact of missense mutations of proteins (D2HGDH and IDH2) causing D-2-hydroxyglutaric aciduria 2, Metab. Brain Dis. (2018), https://doi.org/10.1007/s11011-018-0278-3.
[35] D. Thirumal Kumar, H.G. Eldous, Z.A. Mahgoub, C. George Priya Doss, H. Zayed, Computational modelling approaches as a potential platform to understand the molecular genetics association between Parkinson's and Gaucher diseases, Metab. Brain Dis. (2018), https://doi.org/10.1007/s11011-018-0286-3.
[36] A. Kumar, F. Pintus, A. Di Petrillo, R. Medda, P. Caria, M.J. Matos, D. Viña, E. Pieroni, F. Delogu, B. Era, G.L. Delogu, A. Fais, Novel 2-pheynlbenzofuran derivatives as selective butyrylcholinesterase inhibitors for Alzheimer's disease, Sci. Rep. 8 (2018) 4424, https://doi.org/10.1038/s41598-018-22747-2.
[37] A. Kumar, F. Delogu, Dynamical footprint of cross-reactivity in a human autoimmune T-cell receptor, Sci. Rep. 7 (2017), https://doi.org/10.1038/srep42496.
[38] D. Thirumal Kumar, C. George Priya Doss, P. Sneha, I.A. Tayubi, R. Siva, C. Chakraborty, R. Magesh, Influence of V54M mutation in giant muscle protein titin: a computational screening and molecular dynamics approach, J. Biomol. Struct. Dyn. 35 (2017) 917–928, https://doi.org/10.1080/07391102.2016.1166456.
[39] S. P, T.K. D, G.P.D. C, S. R, H. Zayed, Determining the role of missense mutations in the POU domain of HNF1A that reduce the DNA-binding affinity: a computational approach, PLoS One 12 (2017) e0174953, , https://doi.org/10.1371/journal.pone.0174953.
[40] A. Agrahari, C. George Priya Doss, Impact of I30T and I30M substitution in MPZ gene associated with Dejerine-Sottas syndrome type B (DSSB): a molecular modeling and dynamics, J. Theor. Biol. 382 (2015) 23–33, https://doi.org/10.1016/j.jtbi.2015.06.019.
[41] A.K. Agrahari, A. Kumar, R. Siva, H. Zayed, C. George Priya Doss, Substitution impact of highly conserved arginine residue at position 75 in GJB1 gene in association with X-linked Charcot–Marie-tooth disease: a computational study, J. Theor. Biol. 437 (2018) 305–317, https://doi.org/10.1016/j.jtbi.2017.10.028.
[42] A.K. Agrahari, P. Sneha, C. George Priya Doss, R. Siva, H. Zayed, A profound computational study to prioritize the disease-causing mutations in PRPS1 gene, Metab. Brain Dis. (2017) 1–12, https://doi.org/10.1007/s11011-017-0121-2.
[43] A.K. Agrahari, M. Muskan, C. George Priya Doss, R. Siva, H. Zayed, Computational insights of K1444N substitution in GAP-related domain of NF1 gene associated with neurofibromatosis type 1 disease: a molecular modeling and dynamics approach, Metab. Brain Dis. (2018), https://doi.org/10.1007/s11011-018-0251-1.
[44] S. P, E.A. Ebrahimi, S.A. Ghazala, T.K. D, S. R, G. Priya Doss C, H. Zayed, Structural analysis of missense mutations in galactokinase 1 (GALK1) leading to galactosemia type-2, J. Cell. Biochem. (2018), https://doi.org/10.1002/jcb.27097.
[45] A.K. Agrahari, A. Kumar, R. Siva, H. Zayed, C. George Priya Doss, Substitution impact of highly conserved arginine residue at position 75 in GJB1 gene in association with X-linked Charcot–Marie-tooth disease: a computational study, J. Theor. Biol. 437 (2018), https://doi.org/10.1016/j.jtbi.2017.10.028.
[46] C. George Priya Doss, H. Zayed, Comparative computational assessment of the pathogenicity of mutations in the Aspartoacylase enzyme, Metab. Brain Dis. 32 (2017) 2105–2118, https://doi.org/10.1007/s11011-017-0090-5.
[47] A. Mosaeilhy, M.M. Mohamed, C. George Priya Doss, H.S.A. El Abd, R. Gamal, O.K. Zaki, H. Zayed, Genotype-phenotype correlation in 18 Egyptian patients with glutaric acidemia type I, Metab. Brain Dis. 32 (2017) 1417–1426, https://doi.org/10.1007/s11011-017-0006-4.
[48] P. Sneha, D. Kumar Thirumal, H. Tanwar, R. Siva, C. George Priya Doss, H. Zayed, Structural analysis of G1691S variant in the human filamin B gene responsible for larsen syndrome: a comparative computational approach, J. Cell. Biochem. (2017), https://doi.org/10.1002/jcb.25920.
[49] C.G.P. Doss, D.R. Alasmar, R.I. Bux, P. Sneha, F.D. Bakhsh, I. Al-Azwani, R. El Bekay, H. Zayed, Genetic epidemiology of glucose-6-dehydrogenase deficiency in the arab world, Sci. Rep. 6 (2016), https://doi.org/10.1038/srep37284.
[50] S.K. Ali, P. Sneha, J. Priyadharshini Christy, H. Zayed, C. George Priya Doss, Molecular dynamics-based analyses of the structural instability and secondary structure of the fibrinogen gamma chain protein with the D356V mutation, J. Biomol. Struct. Dyn. 35 (2017) 2714–2724, https://doi.org/10.1080/07391102.2016.1229634.
[51] A. Bairoch, R. Apweiler, The SWISS-PROT protein sequence data bank and its new supplement TREMBL, Nucleic Acids Res. 24 (1996) 21–25, https://doi.org/10.1093/nar/24.1.21.
[52] S.T. Sherry, dbSNP: the NCBI database of genetic variation, Nucleic Acids Res. 29 (2001) 308–311, https://doi.org/10.1093/nar/29.1.308.
[53] H.M. Berman, T. Battistuz, T.N. Bhat, W.F. Bluhm, P.E. Bourne, K. Burkhardt, Z. Feng, G.L. Gilliland, L. Iype, S. Jain, P. Fagan, J. Marvin, D. Padilla, V. Ravichandran, B. Schneider, N. Thanki, H. Weissig, J.D. Westbrook, C. Zardecki, The protein data bank, Acta Crystallogr. Sect. D Biol. Crystallogr. 58 (2002) 899–907, https://doi.org/10.1107/S0907444902003451.
[54] A. Bayat, A.W. Juul, Mild Lesch-Nyhan disease in a boy with a null-mutation in HPRT1: an exception to the known genotype-phenotype correlation: three-year follow up, Med. Res. Arch. 4 (2016) 135–137, https://doi.org/10.1007/8904_2014_368.
[55] P.C. Ng, S. Henikoff, SIFT: predicting amino acid changes that affect protein function, Nucleic Acids Res. 31 (2003) 3812–3814, https://doi.org/10.1093/nar/gkg509.
[56] Y. Bromberg, B. Rost, SNAP, Predict effect of non-synonymous polymorphisms on function, Nucleic Acids Res. 35 (2007) 3823–3835, https://doi.org/10.1093/nar/gkm238.
[57] H.A. Shihab, J. Gough, D.N. Cooper, P.D. Stenson, G.L.A. Barker, K.J. Edwards, I.N.M. Day, T.R. Gaunt, Predicting the functional, molecular, and phenotypic consequences of amino acid substitutions using hidden Markov models, Hum. Mutat. 34 (2013) 57–65, https://doi.org/10.1002/humu.22225.
[58] I.A. Adzhubei, S. Schmidt, L. Peshkin, V.E. Ramensky, A. Gerasimova, P. Bork, A.S. Kondrashov, S.R. Sunyaev, A method and server for predicting damaging missense mutations, Nat. Methods 7 (2010) 248–249, https://doi.org/10.1038/nmeth0410-248.
[59] E. Capriotti, P. Fariselli, I. Rossi, R. Casadio, A three-state prediction of single point mutations on protein stability changes, BMC Bioinf. 9 (2008), https://doi.org/10.1186/1471-2105-9-S2-S6.
[60] C.W. Chen, J. Lin, Y.W. Chu, iStable: off-the-shelf predictor integration for predicting protein stability changes, BMC Bioinf. 14 (2013), https://doi.org/10.1186/1471-2105-14-S2-S5.
[61] H. Venselaar, T.A. TeBeek, R.K. Kuipers, M.L. Hekkelman, G. Vriend, Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces, BMC Bioinf. 11 (2010) 548, https://doi.org/10.1186/1471-2105-11-548.
[62] N. Guex, M.C. Peitsch, T. Schwede, Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: a historical perspective, Electrophoresis 30 (2009), https://doi.org/10.1002/elps.200900140.
[63] M.J. Abraham, T. Murtola, R. Schulz, S. Páll, J.C. Smith, B. Hess, E. Lindah, Gromacs: high performance molecular simulations through multi-level parallelism from laptops to supercomputers, SoftwareX 1–2 (2015) 19–25, https://doi.org/10.1016/j.softx.2015.06.001.
[64] A.S. Moffett, K.W. Bender, S.C. Huber, D. Shukla, Molecular dynamics simulations reveal the conformational dynamics of arabidopsis thaliana bri1 and bak1 receptor-like kinases, J. Biol. Chem. 292 (2017) 12643–12652, https://doi.org/10.1074/jbc.M117.792762.
[65] S. Jahandideh, M. Jamalan, M. Faridounnia, Molecular dynamics study of the dominant-negative E219K polymorphism in human prion protein, J. Biomol. Struct. Dyn. 33 (2015) 1315–1325, https://doi.org/10.1080/07391102.2014.945486.
[66] N. Sun, C. Yuan, X. Ma, Y. Wang, X. Gu, W. Fu, Molecular mechanism of action of RORγt agonists and inverse agonists: insights from molecular dynamics simulation, Molecules 23 (2018) 3181, https://doi.org/10.3390/molecules23123181.
[67] M.Y. Lobanov, N.S. Bogatyreva, O.V. Galzitskaya, Radius of gyration as an indicator of protein structure compactness, Mol. Biol. 42 (2008) 623–628, https://doi.org/10.1134/S0026893308040195.
[68] D.L. Theobald, D.S. Wuttke, Accurate structural correlations from maximum likelihood superpositions, PLoS Comput. Biol. 4 (2008), https://doi.org/10.1371/journal.pcbi.0040043.
[69] A. Amadei, A.B.M. Linssen, H.J.C. Berendsen, Essential dynamics of proteins, Proteins Struct. Funct. Bioinforma. 17 (1993) 412–425, https://doi.org/10.1002/prot.340170408.
[70] W. Kabsch, C. Sander, Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features, Biopolymers 22 (1983) 2577–2637, https://doi.org/10.1002/bip.360221211.
[71] Y. Bian, J. Zhang, J. Wang, J. Wang, W. Wang, Free energy landscape and multiple folding pathways of an H-Type RNA pseudoknot, PLoS One 10 (2015), https://doi.org/10.1371/journal.pone.0129089.
[72] E. Alm, D. Baker, Prediction of protein-folding mechanisms from free-energy landscapes derived from native structures, Proc. Natl. Acad. Sci. U. S. A 96 (1999) 11305–11310, https://doi.org/10.1073/PNAS.96.20.11305.