


RESEARCH

Open Access



A comprehensive in silico exploration of the impacts of missense variants on two different conformations of human pirin protein

Auroni Semonti Khan^{1†}, Nahid Parvez^{1†}, Tamim Ahsan^{2*†} , Sabrina Samad Shoily³ and Abu Ashfaqur Sajib³

Abstract

Background: Pirin, a member of the cupin superfamily, is an iron-binding non-heme protein. It acts as a coregulator of several transcription factors, especially the members of NFκB transcription factor family. Based on the redox state of its iron cofactor, it can assume two different conformations and thereby act as a redox sensor inside the nucleus. Previous studies suggested that pirin may be associated with cancer, inflammatory diseases as well as COVID-19 severities. Hence, it is important to explore the pathogenicity of its missense variants. In this study, we used a number of in silico tools to investigate the effects of missense variants of pirin on its structure, stability, metal cofactor binding affinity and interactions with partner proteins. In addition, we used protein dynamics simulation to elucidate the effects of selected variants on its dynamics. Furthermore, we calculated the frequencies of haplotypes containing pirin missense variants across five major super-populations (African, Admixed American, East Asian, European and South Asian).

Results: Among a total of 153 missense variants of pirin, 45 were uniformly predicted to be pathogenic. Of these, seven variants can be considered for further experimental studies. Variants R59P and L116P were predicted to significantly destabilize and damage pirin structure, substantially reduce its affinity to its binding partners and alter pirin residue fluctuation profile via changing the flexibility of several key residues. Additionally, variants R59Q, F78V, G98D, V151D and L220P were found to impact pirin structure and function in multiple ways. As no haplotype was identified to be harboring more than one missense variant, further interrogation of the individual effects of these seven missense variants is highly recommended.

Conclusions: Pirin is involved in the transcriptional regulation of several genes and can play an important role in inflammatory responses. The variants predicted to be pathogenic in this study may thus contribute to a better understanding of the underlying molecular mechanisms of various inflammatory diseases. Future studies should be focused on clarifying if any of these variants can be used as disease biomarkers.

Keywords: Pirin, Pathogenic variants, Non-heme protein, Inflammation, Transcriptional regulation, Oxidative stress, Cancer, NFκB pathway

Background

Pirin (PIR) is a highly conserved protein across both eukaryotes and prokaryotes (Dunwell *et al.* 2001). It is widely present in sub-nuclear structures of cells (Pang *et al.* 2004). This protein is predominantly expressed in human liver, heart, kidneys, and muscles while low levels of expression are observed in brain and lungs (Wendler *et al.* 1997; Pang *et al.* 2004). Human pirin has been implicated in skin, breast, lung, head and neck,

[†]Auroni Semonti Khan, Nahid Parvez and Tamim Ahsan have contributed equally

*Correspondence: tamimahsan9@gmail.com; tamim.ahsan@nib.gov.bd

² Molecular Biotechnology Division, National Institute of Biotechnology, Savar, Dhaka 1349, Bangladesh

Full list of author information is available at the end of the article

gastrointestinal, cervical cancers (Perez-Dominguez et al. 2021) and tumors of epithelial tissues, hematopoietic and neurological systems (Yoshikawa et al. 2004; Licciulli et al. 2010a; Jungk et al. 2016).

The human pirin is a non-heme iron-binding protein that comprises 290 amino acids and has a molecular weight of 32 kDa. Due to its primary sequence and structural similarity, pirin belongs to the cupin superfamily. It contains two structurally similar β -barrel domains facing each other (Wendler et al. 1997; Pang et al. 2004). The N-terminal domain is highly conserved in mammals, plants, fungi and prokaryotic organisms (Wendler et al. 1997), whereas the C-terminal domain varies. The C-terminal domain does not contain any metal-binding site, but has a cavity which is more compact than that of the N-terminal domain (Pang et al. 2004). The iron cofactor (Fe^{2+} or Fe^{3+}) is coordinated by residues His-56, His-58, His-101 and Glu-103 within the negatively charged metal-binding cavity of the N-terminal domain (Pang et al. 2004; Liu et al. 2013). Depending on the redox state, this iron center within the R-shaped surface area allosterically controls the interactions between pirin and its binding partners (Pang et al. 2004; Liu et al. 2013).

Pirin was originally identified as an interactor of nuclear factor I/CCAAT box transcription factor (NFI/CTF1), which can regulate transcription initiation and DNA replication (Santoro et al. 1988; Wendler et al. 1997). Nuclear factor κB (NF- κB), a family of transcription factors, plays crucial roles in intracellular signaling for immune responses and consists of dimeric RelA (p65), RelB, c-Rel, p50, and p52 (Ghosh et al. 1998; Li and Verma 2002). Pirin functions as a transcriptional coregulator for the expression of NF- κB target genes through regulating the binding of NF- κB p65 to κB -sites (Liu et al. 2013). Nuclear translocation of NF- κB and its binding to DNA is prevented by members of the inhibitor of kappa-B ($\text{I}\kappa\text{B}$) family (Bhatt and Ghosh 2014). B-cell lymphoma 3-encoded protein (BCL3), despite being a member of the $\text{I}\kappa\text{B}$ family, has both transactivation and transrepression roles in regulation of p50 or p52 (members of NF- κB family) homodimers mediated pathways (Lenardo and Siebenlist 1994; Maldonado and Melendez-Zajgla 2011). The difference in mechanism of inhibition by $\text{I}\kappa\text{B}$ and stimulation by BCL3 is not known, but it is evident that these bind to different protein partners. Pirin is one of the four binding partners of BCL3 and is known to enhance the DNA binding by BCL3-p50 via formation of a quaternary complex (Dechend et al. 1999).

The nuclear factor erythroid-derived 2-like 2 (Nrf2) transcription factor triggered in response to cellular oxidative stress conditions can modulate *PIR* gene expression via the functional antioxidant response elements (AREs) in its promoter region (Hübner et al. 2009;

Chorley et al. 2012). Chronic cigarette smoking causes increased oxidative stress, which up-regulates expression of pirin and it has been implicated in bronchial epithelial cell apoptosis (Hübner et al. 2009; Chorley et al. 2012). Pirin is an oxidative stress sensor (Perez-Dominguez et al. 2021) and ferric conformation of this protein facilitates binding of NF- κB p65 to DNA (Liu et al. 2013). Although elevated activation of p65 is observed in chronic inflammation (Giridharan and Srinivasan 2018), canonical NF- κB pathway mediated by p65/p50 heterodimer is critically important for innate immunity and inflammatory responses (Oeckinghaus and Ghosh 2009; Liu et al. 2017). Furthermore, a recent study suggested possible association of pirin with the severity of COVID-19 (Shoily et al. 2021).

Pirin may be functionally important and associated with several diseases (Licciulli et al. 2010b; Perez-Dominguez et al. 2021). For example, terminal myeloid differentiation is impaired by down-regulation of *PIR*, and reduction in pirin activity may be involved in differentiation arrest characteristic of acute myeloid leukemia (AML) (Licciulli et al. 2010a). Additionally, loss of pirin function may lead to immunodeficiency by hampering binding of p65 to DNA. Therefore, identification of the pathogenic variants in this protein can help in determining potential genetic factors leading to immunodeficiency and AML as well as other pathological conditions resulting for disruption of pirin function. Missense variants involving amino acid substitutions at functionally important sites may affect protein structure, stability, flexibility, ligand binding and protein-protein interactions (Zhang et al. 2012; Tamura et al. 2017; Hernandez and Facelli 2021; Birolo et al. 2021; Qi et al. 2021). Studying all variants experimentally is time-consuming, and demands significant resources and efforts. In silico analyses can ease the process by prioritizing missense variants for further experimental studies. In the present study, the impacts of pirin missense variants were explored using various in silico tools.

Materials and methods

Retrieval of missense variants of pirin and prediction of pathogenicity of the missense variants

The list of missense variants in pirin was retrieved from the Ensembl Genome Browser (Howe et al. 2021). The amino acid sequence of pirin was retrieved from UniProt Knowledgebase (UniProtKB) (Accession Number: O00625). Five different pathogenicity predicting tools—Sorting Intolerant From Tolerant (SIFT) (Sim et al. 2012), Polymorphism Phenotyping-2 (PolyPhen-2) (Adzhubei et al. 2010), PMut (López-Ferrando et al. 2017), Meta-SNP (Capriotti et al. 2013) and Rhapsody (Ponzoni et al. 2020) were used to predict the pathogenicity of these

variants. Only the variants which were predicted to be harmful by all of these five tools were selected for further analyses.

Determination of the effects of missense variants on the stability of pirin protein

First, 3D models of Fe²⁺ and Fe³⁺ bound conformations of pirin were generated by template-based modeling at SWISS-MODEL server (Waterhouse et al. 2018) using their X-ray crystallographic structures as templates (PDB IDs: 1J1L and 4GUL, respectively). These X-ray crystallographic structures were retrieved from the RCSB Protein Data bank (PDB) (Berman et al. 2000). These 3D structures were used as input in mutation Cutoff Scanning Matrix (mCSM) (Pires et al. 2014), DeepDDG (Cao et al. 2019), Impact of Nonsynonymous variations on Protein Stability 3D (INPS3D) (Savojardo et al. 2016), MAESTROweb (Laimer et al. 2015), and PremPS (Chen et al. 2020) to assess the destabilizing effects of the selected pirin variants.

Structural analysis of pirin variants

The impacts of these variants on pirin structure was analyzed using Missense3D (Ittisoponpisan et al. 2019; Khanna et al. 2021). This is a structure database resource which can predict structural changes due to buried Gly replacement, Gly introduction in a bend, buried Pro introduction, Cys-Pro replacement, buried H-bond breakage, buried charge introduction, buried charge replacement, buried charge switch, buried salt bridge breakage, di-sulfide bond breakage, buried hydrophilic introduction, disallowed phi/psi introduction, cavity alteration, clash and secondary structure alteration. These changes may lead to the disruption of alpha helices or beta sheets, alterations in inter-residue interactions in the structures as well as unfavorable energy changes that affect the protein's structure (Kajander et al. 2000; Betts and Russell 2003; Krieger et al. 2005; Ho and Brasseur 2005; Hubbard and Haider 2010; Chan et al. 2011).

Assessment of the effects of variants on metal cofactor binding of pirin

To evaluate the impacts of the selected variants on pirin-iron interactions, the variant models were analyzed through metal ion-binding site prediction (MIB) (Lin et al. 2016) tool that uses fragment transformation method and can predict the metal-binding sites in a protein as well as perform metal ion docking (Lu et al. 2006).

Determination of effects of missense variants of pirin on protein-protein interactions

The Fe³⁺ conformation of pirin modulates DNA binding by NF- κ B p65 (Liu et al. 2013) and in a previous

study, Ankyrin (ANK) repeat number 5, 6 and 7 of BCL3 was shown to bind to the Fe²⁺ conformation of pirin (Pang et al. 2004). Amino acid sequence of BCL3 covering these repeats (275–367) was retrieved from UniProtKB (Accession no. P20749) (Bateman et al. 2021). X-ray crystallographic structure obtained from RCSB PDB (PDB ID: 1K1A) was used as a template to generate a 3D model of these three repeat regions using SWISS-MODEL (Waterhouse et al. 2018). For p65 (UniProtKB Accession no. Q04207), a 3D model of its Rel homology domain (RHD) (19–306 amino acid residues) was prepared using the template 1RAM (PDB ID). Docking was performed using HDock (Yan et al. 2020) based on previously published docking results (Pang et al. 2004; Liu et al. 2013).

In case of docking with BCL3, the top model was superimposed on BCL3-p50 homodimer structure using UCSF Chimera 1.14 (Pettersen et al. 2021). Previously described method was followed for this superimposition, and our observation was consistent with the published data (Pang et al. 2004). Interactions between PIR-p65 in the top model were evaluated with Protein Interaction Calculator (PIC) web server (Tina et al. 2007). PIR-p65 interactions were consistent with the available information (Liu et al. 2013).

PIR-BCL3 and PIR-p65 complexes thus predicted were then utilized to determine the impacts of pirin missense variants on its interactions with partners using mCSM-PPI2 (Rodrigues et al. 2019), MutaBind2 (Zhang et al. 2020), SAAMBE-3D (Pahari et al. 2020) and BeAtMuSiC V1.0 (Dehouck et al. 2013).

Simulation of the dynamics of pirin variants

To appraise the changes in protein dynamics caused by missense variants, the wild-type and mutant 3D models of pirin protein were subjected to simulation through a CABS coarse-grained protein model (Kmieciak et al. 2016). Previous study has suggested this method to be a proper alternative to conventional molecular dynamics that includes all atoms (Jamroz et al. 2013). Variants that showed significant negative impact in all the aforementioned analyses were selected for the simulation study. 3D models of variant structures were prepared using template-based modeling in SWISS-MODEL after manually substituting the wild-type residues with the mutant ones in the amino acid sequence of pirin. 3D models of both pirin conformations previously prepared with SWISS-MODEL were used as templates (Waterhouse et al. 2018). The dynamics of the wild-type and the variant structures were simulated with CABS-flex 2.0 using its default parameters (Kuriata et al. 2018).

Haplotyping of the *PIR* variants

We were interested to see if any haplotype of *PIR* contains more than one missense variant. For this purpose, all SNPs with global minor allele frequencies > 0.001 were identified from dbSNP (Sherry et al. 2001). Haplotypes containing these polymorphisms and their frequencies were analyzed with the LDhap module of LDlink (Machiela and Chanock 2015). This module was used to calculate haplotype frequencies in five super-populations (African, Admixed American, East Asian, European and South Asian).

Results

Pathogenicity of the missense variants

A total of 153 missense variants of pirin were identified (Additional file 1: Table S1). The pathogenicity of these variants was predicted using five different tools (SIFT, PolyPhen2, PMut, Meta-SNP and Rhapsody). Each tool follows a different algorithm to predict pathogenicity. SIFT uses sequence homology-based approach for classifying a missense variant as either deleterious or tolerated (Sim et al. 2012). PolyPhen-2 classifies an amino acid substitution into probably damaging, possibly damaging, and benign on the basis of sequence, phylogenetic and structural characteristics of the substitution (Adzhubei et al. 2010, 2013). PMut uses neural networks to predict structure and evolutionary properties resulting from change in amino acid sequence (López-Ferrando et al. 2017). Rhapsody utilizes sequence coevolution data along with structure- and dynamics-based methods to predict pathogenicity of target variants (Ponzoni et al. 2020). Meta-SNP discriminates between disease-related and polymorphic nonsynonymous SNVs (nsSNV) through a random forest-based binary classification method and utilizes several different prediction tools to derive a consensus result (Capriotti et al. 2013). In this study, 45 variants were predicted to have harmful effects by all tools (Table 1). These variants were selected for further analyses.

Destabilizing effects of the potentially pathogenic pirin variants

We used five different tools (mCSM, DeepDDG, INPS3D, MAESTROweb and PremPS) to evaluate the impacts of the variants on pirin stability. mCSM exploits the correlation between the impact of a mutation and atomic distance patterns surrounding the amino acid residue using graph-based signatures to predict stability changes (Pires et al. 2014). DeepDDG relies on neural network-based methods in the prediction of changes in protein stability due to point mutations (Cao et al. 2019). INPS3D uses descriptors to calculate $\Delta\Delta G$ values using a support vector regression (Savojardo et al. 2016). Descriptors

extracted from the protein sequence to differentiate between wild and changed protein include a substitution score derived from the Blosum62 matrix, Kyte–Doolittle hydrophobicity scores of native and changed, the mutability index of the native residue, the molecular weights of native and changed residues, the difference in the alignment score between the native and variant sequences and an HMM, encoding evolutionary information of the wild-type sequence. MAESTROweb utilizes a multi-agent machine learning system based on protein structure to produce changes in unfolding free energy upon point mutation (Laimer et al. 2015). PremPS uses random forest regression scoring function to estimate effects of single mutations on protein stability. It employs an energy function to calculate unfolding Gibbs free energy (Chen et al. 2020).

The predicted changes in folding Gibbs free energy ($\Delta\Delta G$) of both ferrous and ferric Pirin by these variants are listed in Table 2. $\Delta\Delta G$ value < -1.0 kcal/mol (shown as bold) was considered to reflect destabilization and variants predicted by at least four tools to have $\Delta\Delta G$ value < -1.0 kcal/mol were considered to have significant destabilizing effects. Based on this criterion, nine variants identified to have significant destabilizing impacts on both conformations, whereas G98D may be more destabilizing for Fe^{2+} conformation as compared to the Fe^{3+} one.

Structural changes caused by the potentially pathogenic variants

Missense3D was used to predict the effects of the missense variants on pirin structure. First, 21 variants were predicted to have structural damages in the Fe^{2+} bound state. For the Fe^{3+} bound conformation, 20 variants were predicted to alter pirin conformation (Table 3). Six of the missense variants with bound Fe^{2+} (D43H, R59P, G60V, F78V, H101Y and D173G) were predicted to cause buried hydrogen breakage (Additional file 2: Table S2). Variants H101Y and D173G were predicted to alter cavity volume by more than 70 \AA^3 in both the Fe^{2+} and Fe^{3+} bound states. Besides, H58R and G83D were predicted to have altered cavities only in the Fe^{3+} bound conformation (Table 3).

Effects of variants on metal cofactor binding

Since the iron-binding site of pirin is crucial for its biological activity, the MIB tool (Lin et al. 2016) was used to investigate the effects of missense variants of pirin on iron binding. This tool uses fragment transformation method for binding site prediction and docking of metal ions. The overall binding scores of Fe^{2+} to pirin protein were similar for wild-type structure and the missense variant, except H56Q, H58R and H101Y (Additional

Table 1 Pathogenic missense variants of pirin protein as predicted by five separate tools

Variant ID	Variants	SIFT	PolyPhen-2 ^a	PMut	Meta-SNP	Rhapsody
rs372506134	G19A	Deleterious	Probably damaging	Disease	Disease	Deleterious
rs758349788	V24F	Deleterious	Probably damaging	Disease	Disease	Deleterious
rs1485928589	R25W	Deleterious	Possibly damaging	Disease	Disease	Deleterious
rs766252248	I28T	Deleterious	Probably damaging	Disease	Disease	Deleterious
rs746253345	P38L	Deleterious	Probably damaging	Disease	Disease	Deleterious
rs1277921319	D43H	Deleterious	Probably damaging	Disease	Disease	Deleterious
rs1356176104	H56Q	Deleterious	Probably damaging	Disease	Disease	Deleterious
rs780168534	H58R	Deleterious	Probably damaging	Disease	Disease	Deleterious
rs188288097	R59P	Deleterious	Probably damaging	Disease	Disease	Deleterious
rs188288097	R59Q	Deleterious	Probably damaging	Disease	Disease	Deleterious
rs1319331957	G60S	Deleterious	Probably damaging	Disease	Disease	Deleterious
rs1459166472	G60V	Deleterious	Probably damaging	Disease	Disease	Deleterious
rs780078643	G70A	Deleterious	Probably damaging	Disease	Disease	Deleterious
rs748257098	G70R	Deleterious	Probably damaging	Disease	Disease	Deleterious
rs780078643	G70V	Deleterious	Probably damaging	Disease	Disease	Deleterious
rs750390136	D77E	Deleterious	Probably damaging	Disease	Disease	Deleterious
rs1042818236	F78V	Deleterious	Probably damaging	Disease	Disease	Deleterious
rs757045955	H81P	Deleterious	Probably damaging	Disease	Disease	Deleterious
rs1464579620	G83D	Deleterious	Probably damaging	Disease	Disease	Deleterious
rs866898423	L90F	Deleterious	Probably damaging	Disease	Disease	Deleterious
rs149497039	A95V	Deleterious	Probably damaging	Disease	Disease	Deleterious
rs778749014	G98D	Deleterious	Probably damaging	Disease	Disease	Deleterious
rs140109164	G98S	Deleterious	Probably damaging	Disease	Disease	Deleterious
rs1467570812	H101Y	Deleterious	Probably damaging	Disease	Disease	Deleterious
rs752062795	Q115K	Deleterious	Probably damaging	Disease	Disease	Deleterious
rs1484554733	L116P	Deleterious	Probably damaging	Disease	Disease	Deleterious
rs1329364366	M126T	Deleterious	Probably damaging	Disease	Disease	Deleterious
rs1996173	P129L	Deleterious	Probably damaging	Disease	Disease	Deleterious
rs1294033379	V151D	Deleterious	Probably damaging	Disease	Disease	Deleterious
rs772251328	S161Y	Deleterious	Probably damaging	Disease	Disease	Deleterious
rs953093600	T167I	Deleterious	Possibly damaging	Disease	Disease	Deleterious
rs768193675	D173G	Deleterious	Probably damaging	Disease	Disease	Deleterious
rs780763035	D173N	Deleterious	Probably damaging	Disease	Disease	Deleterious
rs779413343	G179V	Deleterious	Possibly damaging	Disease	Disease	Deleterious
rs760795372	P187L	Deleterious	Possibly damaging	Disease	Disease	Prob.delet
rs1569195774	W190S	Deleterious	Possibly damaging	Disease	Disease	Deleterious
rs751833973	L220P	Deleterious	Probably damaging	Disease	Disease	Deleterious
rs1272804008	P245S	Deleterious	Probably damaging	Disease	Disease	Prob.delet
rs772771810	E248A	Deleterious	Probably damaging	Disease	Disease	Deleterious
rs769242287	E248D	Deleterious	Probably damaging	Disease	Disease	Deleterious
rs747391287	G254C	Deleterious	Probably damaging	Disease	Disease	Deleterious
rs996737505	G254V	Deleterious	Probably damaging	Disease	Disease	Deleterious
rs762648888	V257A	Deleterious	Probably damaging	Disease	Disease	Deleterious
rs764770692	M258I	Deleterious	Probably damaging	Disease	Disease	Deleterious
rs761242213	I264S	Deleterious	Probably damaging	Disease	Disease	Deleterious

^a PolyPhen-2 classifies variants as benign, possibly damaging and probably damaging. Only those variants that were predicted to be "probably damaging" were considered to be harmful to increase the accuracy of prediction

Table 2 Impacts of potentially pathogenic missense variants on pirin stability

Variants ^a	Changes in Fe ²⁺ bound Pirin Stability, $\Delta\Delta G$ (kcal/mol) ^b					Variants ^a	Changes in Fe ³⁺ bound Pirin Stability, $\Delta\Delta G$ (kcal/mol) ^b				
	mCSM	DeepDDG	INPS3D	MAESTROweb	PremPS		mCSM	DeepDDG	INPS3D	MAESTROweb	PremPS
G19A	-0.281	-1.309	0.12	-0.045	-0.85	G19A	-0.348	-1.419	0.099	-0.036	-0.97
V24F	-1.211	-0.729	-0.38	0.736	-1.19	V24F	-1.27	-1.047	-0.395	0.723	-1.35
R25W	-0.567	-0.445	-0.574	0.396	-1.27	R25W	-0.577	-0.109	-0.276	0.348	-1.24
I28T	-2.411	-1.377	-2.543	-1.268	-2.03	I28T	-3.068	-1.927	-2.511	-1.317	-2.4
P38L	-0.932	-3.204	-0.523	0.592	-0.73	P38L	-0.965	-3.391	-0.523	0.426	-0.73
D43H	-1.98	-0.887	-0.623	0.121	-0.55	D43H	-1.564	-0.509	-0.679	0.212	-0.71
H56Q	-0.754	-0.396	-0.853	1	-1.33	H56Q	-0.9	-0.529	-0.853	0.246	-1.54
H58R	-1.891	-1.472	-0.258	-0.005	-1.51	H58R	-1.97	-1.444	-0.321	-0.016	-1.59
R59P	-1.594	-2.737	-1.567	-1.498	-1.41	R59P	-1.571	-2.829	-1.612	-1.767	-1.56
R59Q	-1.478	-1.578	-1.013	-1.66	-1.94	R59Q	-1.485	-1.538	-1.068	-1.781	-2.09
G60S	-1.841	-3.177	-0.912	-0.479	-1.56	G60S	-1.889	-3.076	-0.912	-0.531	-1.69
G60V	-0.605	-3.761	-1.185	0.065	-1.44	G60V	-0.6	-3.656	-1.185	0.252	-1.52
G70A	-0.78	-2.001	-0.099	-0.102	-0.3	G70A	-0.811	-1.701	-0.125	-0.014	-0.4
G70R	-0.746	-1.566	0.119	-0.045	-0.53	G70R	-0.823	-1.287	0.227	-0.027	-0.62
G70V	-0.685	-2.375	0.181	0.557	-0.44	G70V	-0.693	-2.049	0.21	0.218	-0.52
D77E	-0.601	-1.97	-0.586	0.319	-0.7	D77E	-0.422	-2.908	-0.65	0.328	-0.77
F78V	-1.476	-3.941	-2.152	-0.739	-1.95	F78V	-1.549	-4.009	-2.086	-1.03	-2.23
H81P	0.725	-1.264	-1.286	0.019	-0.95	H81P	0.696	-1.189	-1.242	0.069	-1.13
G83D	-1.452	-1.512	-0.683	0.599	-1.27	G83D	-1.473	-1.529	-0.698	0.697	-1.45
L90F	-1.604	-1.029	-1.032	0.437	-1.02	L90F	-1.585	-0.804	-1.011	0.441	-1.16
A95V	-0.963	-2.69	-1.356	<u>1.485</u>	-1.81	A95V	-0.899	-2.24	-1.356	<u>1.541</u>	-1.9
G98D	-1.632	-3.5	-0.963	-1.203	-1.64	G98D	-1.789	-3.266	-0.987	-0.548	-1.78
G98S	-1.299	-2.466	-0.93	0.106	-1.31	G98S	-1.421	-2.281	-0.938	0.182	-1.45
H101Y	-0.182	-0.845	-0.116	0.454	-1.02	H101Y	-0.016	-0.744	-0.116	0.44	-1.08
Q115K	-0.339	-1.437	-0.796	-0.556	-1.46	Q115K	-0.263	-1.195	-0.607	-0.355	-1.34
L116P	-1.323	-4.818	-3.54	-1.508	-2.79	L116P	-1.437	-4.762	-3.54	-1.645	-3.02
M126T	-0.114	-1.028	-1.544	-1.123	-1.37	M126T	-0.145	-0.798	-1.533	-1.237	-1.27
P129L	-0.713	-1.614	-0.579	0.436	-0.69	P129L	-0.748	-1.719	-0.579	0.303	-0.85
V151D	-3.091	-5.314	-3.003	-1.201	-2.79	V151D	-3.117	-5.492	-3.003	-1.18	-2.94
S161Y	-0.62	-1.719	-0.524	0.85	-0.84	S161Y	-0.607	-1.933	-0.551	0.947	-0.89
T167I	-0.191	-0.479	-0.846	0.177	-1.21	T167I	-0.158	-0.518	-0.846	0.512	-1.21
D173G	0.64	-2.656	-0.949	-0.026	0.01	D173G	0.482	-2.415	-1.005	-0.021	-0.09
D173N	-0.137	-1.544	-0.597	0.064	-0.5	D173N	-0.336	-1.255	-0.742	0.141	-0.48
G179V	-0.612	-2.029	-1.538	-0.93	-1.44	G179V	-0.56	-2.096	-1.542	-0.361	-1.52
P187L	-0.369	-0.886	-0.814	0.066	-0.78	P187L	-0.343	-0.91	-0.814	0.037	-0.85
W190S	-2.767	-1.464	-2.712	-1.465	-1.26	W190S	-2.707	-1.415	-2.654	-1.568	-1.42
L220P	-1.386	-5.855	-3.385	-1.341	-2.32	L220P	-1.379	-5.872	-3.402	-1.261	-2.49
P245S	-1.457	-1.37	-1.046	-0.413	-1.24	P245S	-1.554	-1.494	-1.046	-0.419	-1.45
E248A	-0.689	-0.603	-0.581	-1.544	-0.51	E248A	-0.645	-0.614	-0.575	-1.524	-0.67
E248D	-0.685	-0.853	-0.729	-0.862	-1.03	E248D	-0.597	-0.835	-0.704	-0.879	-1.19
G254C	-0.606	-0.619	-1.685	-0.543	-1.18	G254C	-0.511	-0.649	-1.636	-0.435	-1.2
G254V	-0.199	-0.862	-1.482	-0.709	-1.56	G254V	-0.149	-0.94	-1.4	-0.623	-1.4
V257A	-2.006	-1.87	-2.659	-1.145	-2.1	V257A	-2.155	-1.948	-2.671	-1.143	-2.24
M258I	-0.971	-1.677	-2.049	0.268	-1.67	M258I	-0.921	-1.283	-2.079	0.352	-1.82
I264S	-2.812	-2.587	-3.226	-1.715	-2.56	I264S	-2.676	-2.499	-3.258	-1.757	-2.69

^a Variants that were predicted to significantly reduce pirin stability ($\Delta\Delta G < -1$ kcal/mol) by all five tools are written in bold italics, whereas those that were predicted to have $\Delta\Delta G < -1$ kcal/mol by four tools are written in bold letters

^b Negative $\Delta\Delta G$ values indicate decrease in pirin stability. $\Delta\Delta G$ values of < -1 kcal/mol are written in bold letters, while $\Delta\Delta G$ values of > 1 kcal/mol are underlined

Table 3 Effects of potentially pathogenic missense variants on pirin structure

Variant ^a	Fe2 + conformation Structural Changes	Variant ^a	Fe3 + conformation Structural Changes
G19A	No structural damage detected	G19A	No structural damage detected
V24F	No structural damage detected	V24F	No structural damage detected
R25W	No structural damage detected	R25W	No structural damage detected
I28T	No structural damage detected	I28T	No structural damage detected
P38L	Cis pro replaced	P38L	This substitution triggers clash alert
D43H	Buried H-bond breakage, buried salt bridge breakage	D43H	Buried charge switch, Buried H-bond breakage, Buried salt bridge breakage
H56Q	No structural damage detected	H56Q	No structural damage detected
H58R	No structural damage detected	H58R	Cavity altered
R59P	Buried Pro-introduced, secondary structure altered, disallowed phi/psi, buried charge replaced, buried H-bond breakage, buried salt bridge breakage	R59P	Buried Pro introduced, secondary structure altered, disallowed phi/psi, buried charge replaced, buried H-bond breakage
R59Q	Buried charge replaced, buried salt bridge breakage	R59Q	Buried charge replaced, buried H-bond breakage
G60S	Buried Gly replaced, Gly in a bend	G60S	Disallowed phi/psi, buried Gly replaced, Gly in a bend
G60V	Buried Gly replaced, buried H-bond breakage, Gly in a bend	G60V	Clash, disallowed phi/psi, buried Gly replaced, buried H-bond breakage, Gly in a bend
G70A	Disallowed phi/psi, Gly in a bend	G70A	Clash, disallowed phi/psi, buried Gly replaced, buried H-bond breakage, Gly in a bend
G70R	Clash, disallowed phi/psi, Gly in a bend	G70R	Clash, disallowed phi/psi, Gly in a bend
G70V	Disallowed phi/psi, Gly in a bend	G70V	Disallowed phi/psi, Gly in a bend
D77E	No structural damage detected	D77E	No structural damage detected
F78V	Buried H-bond breakage	F78V	No structural damage detected
H81P	No structural damage detected	H81P	No structural damage detected
G83D	No structural damage detected	G83D	Cavity altered
L90F	No structural damage detected	L90F	No structural damage detected
A95V	No structural damage detected	A95V	No structural damage detected
G98D	Buried charge introduced, disallowed phi/psi, buried Gly replaced	G98D	Buried charge introduced, disallowed phi/psi, buried Gly replaced
G98S	Disallowed phi/psi, buried Gly replaced	G98S	Disallowed phi/psi, buried Gly replaced
H101Y	Buried charge replaced, buried H-bond breakage, cavity altered	H101Y	Buried charge replaced, buried H-bond breakage, cavity altered
Q115K	No structural damage detected	Q115K	No structural damage detected
L116P	Buried Pro introduced, Disallowed phi/psi	L116P	Buried Pro introduced, disallowed phi/psi
M126T	No structural damage detected	M126T	No structural damage detected
P129L	No structural damage detected	P129L	No structural damage detected
V151D	Buried hydrophilic introduced, buried charge introduced	V151D	Buried hydrophilic introduced, buried charge introduced
S161Y	No structural damage detected	S161Y	No structural damage detected
T167I	No structural damage detected	T167I	No structural damage detected
D173G	Buried H-bond breakage Cavity altered, buried / exposed switch	D173G	Buried H-bond breakage, cavity altered, buried / exposed switch
D173N	Buried charge replaced	D173N	No structural damage detected
G179V	No structural damage detected	G179V	No structural damage detected
P187L	No structural damage detected	P187L	No structural damage detected
W190S	No structural damage detected	W190S	No structural damage detected
L220P	Buried Pro introduced	L220P	Buried Pro introduced
P245S	No structural damage detected	P245S	No structural damage detected
E248A	No structural damage detected	E248A	No structural damage detected
E248D	No structural damage detected	G254C	Disallowed phi/psi
G254C	Disallowed phi/psi	G254V	Disallowed phi/psi
G254V	Disallowed phi/psi	E248D	No structural damage detected
V257A	No structural damage detected	V257A	No structural damage detected
M258I	No structural damage detected	M258I	No structural damage detected
I264S	No structural damage detected	I264S	No structural damage detected

^a Variants that were predicted to be damaging to pirin structure are written in bold letters

file 3: Table S3). This pattern was observed for Fe³⁺, too. Residues His-56, His-58 and His-101 are three of four residues reported in a previous study to be the constituents of metal-binding site (Pang et al. 2004).

Effects of missense variants on protein–protein interactions

After generation of models for BCL3 and p65 through template-based modeling using SWISS-MODEL, pirin was docked to each of these proteins using HDOCK (Fig. 1) (Yan et al. 2020). Interactions of pirin with BCL3 and p65 were then analyzed for the impacts of missense variants in pirin using mCSM-PPI2 (Rodrigues et al. 2019), MutaBind2 (Zhang et al. 2020), SAAMBE-3D (Pahari et al. 2020) and BeAtMuSiC V1.0 (Dehouck et al. 2013). mCSM-PPI2 predicts the effects of missense variants on protein–protein binding affinity by concentrating on the inter-residue non-covalent interaction network using optimized graph-based signatures like graph kernels, evolutionary information, complex network metrics and energetic terms (Rodrigues et al. 2019). MutaBind2 exploits molecular mechanics force fields, statistical potentials and fast side-chain optimization algorithms built via random forest method (Zhang et al. 2020). SAAMBE-3D uses machine learning (Pahari et al. 2020). BeAtMuSiC V1.0 utilizes a set of statistical potentials derived from known protein structures and combines the effect of the mutation on the strength of the interactions at the interface, and on the overall stability of the complex (Dehouck et al. 2013).

In this study, negative values for $\Delta\Delta G$ indicated reduction in protein–protein binding affinity, and variants with predicted $\Delta\Delta G < -1$ kcal/mole (by at least three tools) were considered to significantly decrease such affinity. Four variants (R59P, F78V, H81P and L116P) in case of pirin-BCL3 were found to be substantially destabilizing in case of pirin-BCL3 complex (Table 4). However, no variant had such considerable effect on pirin-p65 interactions. Interestingly, several variants (G98D, P129L, W190S and L220P in case of pir-BCL3 complex, and R59P, L116P and G254C in case of pirin-p65 complex) were predicted to have $\Delta\Delta G < -1$ kcal/mol by two tools and < -0.85 kcal/mol by at least one of the remaining two tools (Table 4). Such variants might also have significant destabilizing effects.

Alterations in the dynamics of pirin

Two variants (R59P and L116P) were found to destabilize pirin (Table 2), damage its structure (Table 3) and diminish pirin's binding affinity to its partners (Table 4). Protein dynamics simulation was performed for these two variant structures, along with the wild-type ones, in both Fe²⁺ and Fe³⁺ bound conformations. Residue fluctuation

profile revealed three areas with major alterations in fluctuation (Fig. 2). These areas are roughly from residue 28 to 37, 78 to 84, 120 to 125. The first and third regions were found to have altered fluctuation for both variants in both conformations. However, the second region, from 78 to 84, showed large alteration for Fe²⁺ bond conformation only. Region of amino acids 28 to 37 overlap the R-shaped region responsible for p65 binding. 78 to 84 and 120 to 125 amino acid residues overlap two of the acidic patches responsible for BCL3 binding (Fig. 1). There are other regions for each of the two variants that show altered fluctuation but the two conserved iron-binding clusters show no major fluctuation in any case.

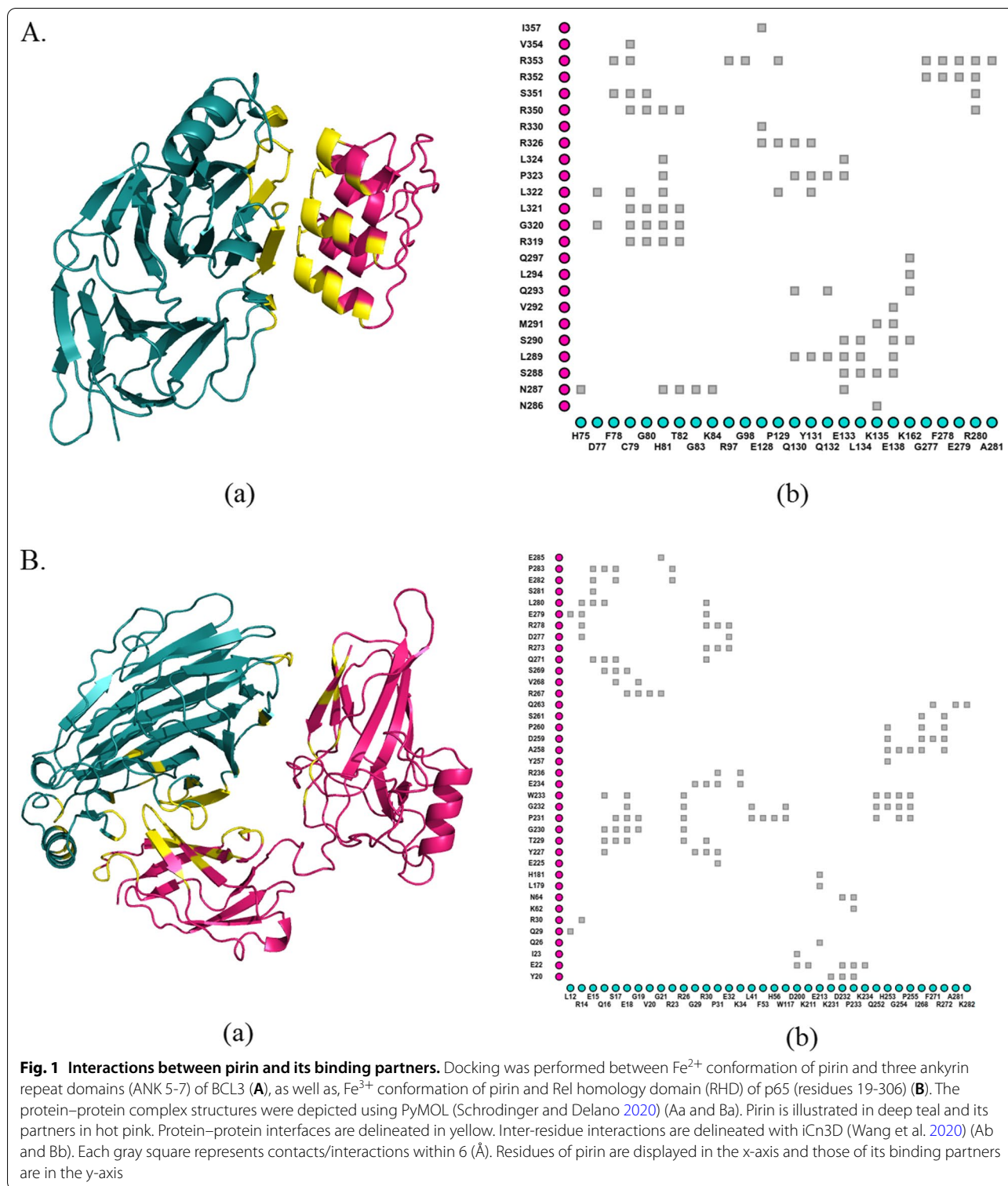
Haplotyping of the variants

With the goal of identifying haplotypes containing more than one pirin missense variants, haplotype frequencies in five super-populations were calculated (Table 5). Three haplotypes were found to harbor two variant alleles. Each of these haplotypes however contained rs8094T allele, which is synonymous.

Discussion

The present study aimed at prioritizing pathogenic missense variants of pirin protein for further experimental analysis. For this purpose, out of 153 missense variants or pirin, 45 were selected for further analyses as they were uniformly predicted to be pathogenic (Table 1). The impacts of these selected variants on pirin's stability (Table 2), structure (Table 3), cofactor binding (Additional file 3: Table S3) and interactions with binding partners (Table 4) were predicted using multiple *in silico* tools to enhance prediction accuracy. Additionally, two variant structures (R59P and L116P) were subjected to protein dynamics simulation (Fig. 2) due to their predicted substantial effects on pirin structure and possibly function. Our findings indicate that these two variants, along with five other variants (R59Q, F78V, G98D, V151D and L220P), should be subjected to further experimental investigations.

Two domains of pirin are the N-terminal domain spanning residues 3–134 and the C-terminal domain comprising residues 143–290 (Pang et al. 2004). Sequence of the N-terminal domain has remained significantly conserved in mammals, plants and prokaryotes, notably in two regions spanning residues 52–70 (cluster 1) and residues 88–106 (cluster 2) (Wendler et al. 1997). Together these regions contain a total of four metal coordinating residues that are strictly conserved among species. N-terminal domain and iron-binding residues are thus important for pirin function. R59P lie in the conserved cluster 1 and L116P is close to conserved cluster 2, both residues located in the N-terminal. The C-terminal domain



contain neither any metal-binding site nor conserved residues (Pang et al. 2004).

Earlier studies reported that pirin interacts with BCL3 and p65 which are members of the IκB and NF-κB family,

respectively (Dechend et al. 1999; Liu et al. 2013). A large acidic patch with residues 77–82, 97–103, and 124–128 is present on the surface of pirin N-terminal domain. This patch has been shown to interact with the large basic

Table 4 Effects of potentially pathogenic missense variants on stability of pirin-BCL3 and pirin-p65 interactions

Variants ^a	Changes in pirin-BCL3 complex stability, $\Delta\Delta G$ (kcal/mol) ^b				Variants ^a	Changes in pirin/p65 complex stability, $\Delta\Delta G$ (kcal/mol) ^b			
	mCSM-PPI2	MutaBind2	SAAMBE3D	BeAtMuSic V1.0		mCSM-PPI2	MutaBind2	SAAMBE3D	BeAtMuSic V1.0
G19A	-0.119	-1	-0.1	-0.25	G19A	-0.881	-0.88	-0.43	-1.29
V24F	0.229	-0.98	0.29	-0.08	V24F	0.407	-0.66	0.03	-0.2
R25W	-0.171	-0.43	-0.34	-0.22	R25W	0.019	-0.26	-0.34	-0.05
I28T	-0.195	-0.74	0.5	-0.84	I28T	-0.19	-0.64	0.5	-0.93
P38L	-0.51	-0.93	0.1	-0.58	P38L	-0.44	-0.57	0.1	-0.41
D43H	-0.11	-0.63	-0.07	0.09	D43H	0.171	-1.09	-0.15	0.15
H56Q	-0.339	-0.47	-0.15	-0.31	H56Q	-0.459	-0.46	-0.38	-0.47
H58R	-0.015	-0.9	-0.18	-0.34	H58R	-0.055	-0.39	-0.29	-0.44
R59P	-1.139	-1.15	-1.9	-1.1	R59P	-1.102	-0.66	-1.9	-0.97
R59Q	-0.656	-0.74	-0.72	-0.65	R59Q	-0.462	-0.3	-0.72	-0.48
G60S	-0.673	-0.99	0.15	-0.6	G60S	-0.438	-0.42	0.15	-0.59
G60V	-0.999	-1.85	-0.16	-0.46	G60V	-0.866	-0.74	-0.16	-0.46
G70A	-0.114	-0.46	0	-0.9	G70A	-0.017	-0.25	0	-0.67
G70R	-0.309	-0.31	-0.34	-0.92	G70R	-0.248	-0.08	-0.34	-0.81
G70V	-0.43	-0.54	-0.16	-0.71	G70V	-0.308	0	-0.16	-0.68
D77E	-0.668	-1.3	-0.7	-0.95	D77E	-0.059	-0.62	-0.6	-0.65
F78V	-1.472	-2.3	-1.16	-0.99	F78V	-0.47	-0.36	-0.85	-0.7
H81P	-2.528	-2.97	-1.57	-1.86	H81P	-0.674	-0.49	-0.79	-0.81
G83D	-0.286	-1.74	-0.61	-0.87	G83D	0.167	-0.49	0.03	-0.43
L90F	0.712	-0.95	-0.03	-0.06	L90F	0.401	-0.73	-0.03	-0.13
A95V	0.12	-0.78	0.43	0.02	A95V	0.207	-0.77	0.43	-0.09
G98D	-0.959	-1.84	0.05	-1.47	G98D	-0.464	-0.54	0.05	-1.3
G98S	-0.938	-0.82	0.04	-1.18	G98S	-0.512	-0.66	0.04	-1.09
H101Y	-0.336	-1.1	-0.72	-0.06	H101Y	-0.077	-0.83	-0.72	-0.09
Q115K	-0.452	-0.49	-0.26	-0.47	Q115K	-0.322	-0.38	-0.26	-0.41
L116P	-1.048	-1.3	-1.11	-1.71	L116P	-0.875	-0.61	-1.11	-1.91
M126T	-0.026	-0.25	-0.35	-0.14	M126T	-0.222	-0.43	-0.35	-0.02
P129L	-1.058	-1.19	0.13	-0.99	P129L	-0.083	-0.48	0.28	-0.1
V151D	-0.418	-0.81	-0.15	-1.05	V151D	-0.364	-0.63	-0.15	-1.17
S161Y	-0.079	-1.26	0.13	-0.18	S161Y	0.142	-0.5	0.13	-0.24
T167I	-0.279	-1.02	0.03	-0.15	T167I	-0.154	-0.43	0.03	-0.09
D173G	-0.5	-1.1	-0.5	-0.34	D173G	-0.38	-0.42	-0.5	-0.38
D173N	-0.608	-0.52	-1.24	-0.05	D173N	-0.792	-0.71	-1.24	-0.13
G179V	-0.02	-0.92	-0.23	-0.61	G179V	-0.075	-0.38	-0.23	-0.57
P187L	-0.248	-0.46	0.13	-0.39	P187L	-0.23	-0.4	0.13	-0.45
W190S	-0.117	-1.19	-1.19	-0.99	W190S	-0.21	-0.51	-1.19	-0.9
L220P	-0.974	-1.22	-0.84	-1.15	L220P	-0.931	-0.45	-0.84	-1.11
P245S	-0.16	-0.88	0.15	-0.5	P245S	-0.131	-0.51	0.15	-0.53
E248A	-0.118	-0.24	-0.32	-0.38	E248A	-0.157	-0.17	-0.32	-0.3
E248D	-0.003	-0.46	-0.51	-0.27	E248D	0.042	-0.21	-0.51	-0.45
G254C	-0.153	-0.61	0.03	-0.48	G254C	-0.897	-1.22	-0.03	-1.08
G254V	-0.079	-0.95	-0.09	-0.72	G254V	-0.62	-0.95	-0.56	-1.55
V257A	-0.316	-0.99	-0.37	-1.04	V257A	-0.324	-0.69	-0.49	-1.14
M258I	-0.357	-0.9	-0.23	-0.07	M258I	-0.387	-1.06	-0.23	-0.15
I264S	-0.301	-0.95	0.08	-0.98	I264S	-0.28	-0.49	0.08	-0.9

^a Variants that were predicted to significantly reduce protein-protein binding affinity ($\Delta\Delta G < -1$ kcal/mol) by at least three tools are written in bold italics. Variants with predicted $\Delta\Delta G < -1$ kcal/mol by two tools and with predicted $\Delta\Delta G$ values between -0.85 and -1 kcal/mol by at least one tool are written in bold letters

^b Negative $\Delta\Delta G$ values indicate decrease in protein-protein binding affinity. $\Delta\Delta G$ values of < -1 kcal/mol are written in bold letters, while $\Delta\Delta G$ values between -0.85 and -1 kcal/mol are written in italics

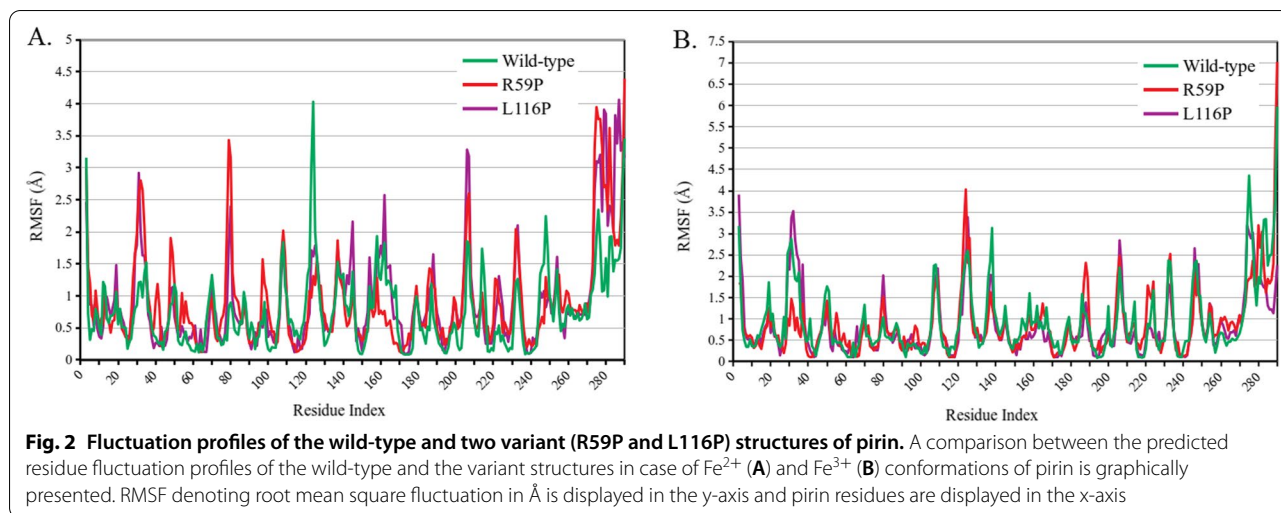


Table 5 Frequencies of haplotypes containing pirin variants

SNP IDs	Haplotypes ^a	Haplotype frequencies					
		ALL	African	American	East Asian	European	South Asian
rs75378219_rs34104000_rs8094_rs35715407_rs34149789	C_A_C_C_C	0.5272	0.5902	0.4828	0.5314	0.4295	0.571
	C_A_T_C_C	0.4286	0.2602	0.4847	0.4686	0.5705	0.429
	C_G_C_C_C	0.0238	0.0788	0.021	–	–	–
	T_A_C_C_C	0.0056	0.0189	0.0038	–	–	–
	C_A_T_C_G	0.0045	0.016	0.0019	–	–	–
	C_A_C_A_C	0.0029	0.011	–	–	–	–
	T_A_T_C_C	0.0029	0.01	0.0019	–	–	–
	C_A_C_C_G	0.0024	0.007	0.0038	–	–	–
C_G_T_C_C	0.0021	0.008	–	–	–	–	

^a Variant DNA bases are written in bold letters, and haplotypes containing more than one variant are written in italics

patch on ankyrin repeats 6 and 7 of BCL3, whereas residues in C-terminal domain may interact with ankyrin repeat 5 (Pang et al. 2004). An R-shaped region is a notable pirin surface region which directly interacts with p65, comprising most of the central binding surface between them (Liu et al. 2013). It is made up of the residues 7–41 and 53–62, which also spans the surrounding area of the metal-binding cavity at the N-terminus (Liu et al. 2013). Arg59 lies in this R-shaped region needed for p65 binding. On the other hand, Leu116 is located in between two acidic patches on the surface of pirin required for interaction with BCL3. Therefore, R59P and L116P are likely to have significant impacts on pirin functionalities.

Residue fluctuation profiles indicated both R59P and L116P cause significant fluctuation in amino acid residues 28 to 37, which is part of the R-shaped region responsible for p65 binding. The two variants also alter fluctuation in regions of 78 to 84 and 120 to 125 amino acid residues which overlap the acidic patches of residues

77 to 82 and 124 to 128 (Fig. 2). Based on previously reported roles of R-shaped region and acidic patch in interaction of pirin with P65 and BCL3 (Pang et al. 2004; Liu et al. 2013), it can be stated that the altered fluctuation due to these variants might hamper the protein–protein interactions of pirin. Our findings are consistent with such assumptions (Table 4).

The metal-binding pocket is situated in the N-terminal domain of pirin and contains residues His-56, His-58, His-101 and Glu-103 which coordinate with ferrous/ferric ion. Metal binding has been reported to be crucial for interactions between pirin and its binding partners. A small-molecule inhibitor TPh A inserts into the Fe²⁺ containing pocket and prevents the cellular activity of pirin by disrupting the formation of pirin-Bcl3 complex (Miyazaki et al. 2010). Moreover, the binding of Fe³⁺ instead of Fe²⁺ alters pirin conformation and thus helps it to bind to p65 of NF-κB family. This iron center plays a role in the allosteric control of

the R-shaped surface loop region (Liu et al. 2013). Pirin may therefore stimulate gene expression by BCL3-p50 complex and p65 in the Fe^{2+} and Fe^{3+} bound states, respectively. Fe^{2+} to Fe^{3+} conversion occurs as the nuclear environment becomes more oxidizing. Hence, the activity of pirin and its related gene expression is dependent on the redox state. The iron-binding center contributes significantly in this process. Since these interactions can be disrupted by the aforementioned missense variants, pathways of the immune system and cell division, modulated by BCL3 and p65 can thus be adversely impacted, leading to disease conditions.

One of the variants that alter the metal-binding residues, H101Y, was found to cause structural damage (Table 3). In addition, the H101Y variant replaces the buried charge, breaks a buried H-bond and alters cavity by 70 Å, in both Fe^{2+} and Fe^{3+} bound conformations (Table 3 and Additional file 2: Table S2). Only in the Fe^{3+} bound state, the variant H58R, another iron coordinating residue, was predicted to cause an alteration in the cavity. In the metal-binding analysis, variants H56Q, H58R and H101Y had scores inconsistent with the wild-type, which may be suggestive of the possible effects of these variants on pirin-iron associations.

L220P in the less conserved C-terminal domain was predicted to alter stability and structure of both Fe^{2+} and Fe^{3+} bound conformations (Table 2 and 3). It may also interfere with interactions with BCL3, but not with p65 (Table 4).

Two other variants of pirin (F78V and G98D) are of considerable interest on the basis of our findings. The F78V variant was found to break buried H-bond in Fe^{2+} bound state only (Additional file 2: Table S2), but it was predicted to destabilize both Fe^{2+} and Fe^{3+} bound conformations. It is part of the acidic patches that bind with ARD of BCL3. In agreement with this information, our study predicted the F78V variant to hamper binding of pirin with BCL3. No such relationship has been observed for p65. In the case of variant G98D, introduction of buried charge and replacement of glycine create structural changes in Fe^{2+} as well as Fe^{3+} bound conformations (Table 3). Destabilization of both conformations and hindrance in interaction with BCL3 is significant for this variant (Tables 2 and 4).

In spite of having no effect on protein-protein interactions, variants R59Q and V151D draw attention as both of these significantly destabilize and alter structures for the Fe^{2+} and Fe^{3+} bound conformations (Table 2 and 3). Four variants (I28T, W190S, V257A and I264S) were identified to significantly destabilize the protein. However, no effect on structure or protein-protein interaction was predicted. In contrast, P38L, D43H, H58R, G60S, G60V, G70A, G70R, G70V, G98S, H101Y, D173G,

D173N, G254C and G254V were found to cause structural damages whereas the stability of those structures was predicted to be unaffected. Here, buried hydrogen bonds were disrupted by D43H, R59P, G60V and D173G (Additional file 2: Table S2). On the other hand, most of the variants displayed similar results regarding redox state of iron, showing damages in both conformations. Strikingly, D173N and H58R were exceptions. Structural damage has been predicted for D173N only when bound to Fe^{2+} and for H58R only when bound to Fe^{3+} . This phenomenon and the differential roles of Asp-173 and His-58 in two pirin conformations should be clarified in future studies.

Although variants P38L, H58R, G60S and G60V are located in the R-shaped region responsible for p65 binding, these residues did not appear to exert any effect on protein-protein interactions (Table 4). The absence of their direct interactions with p65 might explain these results (Fig. 1). His-58 and His-101 are two of the iron coordinating residues and variants at these sites (H58R and H101Y) were found to minimize the binding affinity of iron compared to wild-type protein. According to MIB tool prediction, variants other than the four metal coordinating ones were found to show no adverse effect in iron binding (Additional file 3: Table S3).

Another variant can be of consideration. H81P has been reported by four tools to disrupt interaction with BCL3. Although it has no other negative effect, being part of the acidic patch makes it a significant residue for BCL3 interaction and functionality of pirin.

Absence of any haplotype with more than one missense variant (Table 5) indicates that the presence of multiple missense variants in the same individual is unlikely. So, exploring the combinatorial effect of more than one missense variant in pirin protein may not be necessary, and identifying individual effects of missense variants may be sufficient in this connection.

It should be noted that a recent study identified V257A, I28T and I264S variants to have significant destabilizing effects on pirin structure (Suleman et al. 2021). Our study also found these variants to substantially reduce the stability of both pirin conformations (Table 2). However, the previous study had a lower number of initial missense variants (119, as compared to 153 in our study), and chose less variants for further analysis (24, as compared to 45 in our study). Besides, impacts of variants on pirin structure, cofactor binding and interactions with other proteins were not elaborated. Furthermore, the aforementioned study did not differentiate between the two conformations of pirin. Therefore, our study appears to be the most comprehensive exploration of the effects of pirin missense variants so far. The variants identified in this study for further experimental

clarifications thus contribute to the existing list of prioritized pirin variants.

Conclusions

Since pirin plays a crucial role in regulation of multiple gene expressions, variants that alter its structure and impede its functions can contribute to the pathogenesis of various diseases. Prioritizing these variants for further experimentation is therefore essential. In the present study, we used multiple in silico tools to assess the possible pathogenicity of a total of 153 missense variants and appraise the impacts of a selected set of variants on pirin's structure and functions. Based on our findings, we propose that seven variants (R59P, L116P, L220P, F78V, G98D, R59Q and V151D) should be considered for further investigations. In addition, four other variants (H58R, H101Y, D173N and H81P) can also be important targets of analysis. Since haplotypes with more than one pirin missense variant could not be found, exploring effects of individual variants should be enough for identifying roles of variants in disrupting pirin's functions. Our findings thus significantly contribute to the existing knowledge regarding pathogenic variants of pirin. Future studies should focus on the possibility of using these variants as disease biomarkers.

Abbreviations

NFI/CTF1: Nuclear factor I/CCAAT box transcription factor; NF- κ B: Nuclear factor κ B; I κ B: Inhibitor of kappa-B; BCL3: B-cell lymphoma 3-encoded protein; Nrf2: Nuclear factor erythroid-derived 2-like 2; ARE: Antioxidant response elements; AML: Acute myeloid leukemia; SNP: Single nucleotide polymorphism; nsSNV: Nonsynonymous single nucleotide variants.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s42269-022-00917-7>.

Additional file 1: Table S1. List of missense variants of pirin.

Additional file 2: Table S2. Buried hydrogen bonds broken by pirin missense variants.

Additional file 3: Table S3. Effects of missense variants on metal binding of pirin according to MIB tool.

Acknowledgements

Not applicable.

Author contributions

ASK took part in the acquisition, analysis and interpretation of data, manuscript preparation and reviewing. NP participated in the acquisition, analysis and interpretation of data, manuscript preparation and reviewing. TA was responsible for conception, acquisition, analysis, interpretation of data, manuscript preparation and reviewing. SSS was involved in the analysis, interpretation of data and manuscript reviewing. AAS took part in the analysis, interpretation of data and manuscript reviewing. All authors have read and approved the final manuscript.

Funding

This research was supported by a grant from special allocation in science and technology from the Ministry of Science and Technology, Bangladesh.

Availability of data and materials

All data generated or analyzed during this study are included in this published article (and its Additional files).

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Department of Genetic Engineering and Biotechnology, Jagannath University, Dhaka 1100, Bangladesh. ²Molecular Biotechnology Division, National Institute of Biotechnology, Savar, Dhaka 1349, Bangladesh. ³Department of Genetic Engineering and Biotechnology, University of Dhaka, Dhaka 1000, Bangladesh.

Received: 5 April 2022 Accepted: 21 July 2022

Published online: 30 July 2022

References

- Adzhubei I, Jordan DM, Sunyaev SR (2013) Predicting functional effect of human missense mutations using PolyPhen-2. *Curr Protoc Hum Genet* Chapter 7(Unit7):20. <https://doi.org/10.1002/0471142905.hg0720s76>
- Adzhubei IA, Schmidt S, Peshkin L et al (2010) A method and server for predicting damaging missense mutations. *Nat Methods* 7:248–249. <https://doi.org/10.1038/nmeth0410-248>
- Bateman A, Martin MJ, Orchard S et al (2021) UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res* 49:D480–D489. <https://doi.org/10.1093/NAR/GKAA1100>
- Berman HM, Westbrook J, Feng Z et al (2000) The protein data bank. *Nucleic Acids Res* 28:235–242. <https://doi.org/10.1093/nar/28.1.235>
- Betts MJ, Russell RB (2003) Amino acid properties and consequences of substitutions. *Bioinform Genet* 317:10–1002
- Bhatt D, Ghosh S (2014) Regulation of the NF- κ B-mediated transcription of inflammatory genes. *Front Immunol*. <https://doi.org/10.3389/FIMMU.2014.00071>
- Birolo G, Benevenuta S, Fariselli P et al (2021) Protein stability perturbation contributes to the loss of function in haploinsufficient genes. *Front Mol Biosci* 8:10–15. <https://doi.org/10.3389/fmolb.2021.620793>
- Cao H, Wang J, He L et al (2019) DeepDDG: predicting the stability change of protein point mutations using neural networks. *J Chem Inf Model* 59:1508–1514. https://doi.org/10.1021/ACS.JCIM.8B00697/SUPPL_FILE/C18B00697_SI_003.XLSX
- Capriotti E, Altman RB, Bromberg Y (2013) Collective judgment predicts disease-associated single nucleotide variants. *BMC Genomics* 14(Suppl 3):S2. <https://doi.org/10.1186/1471-2164-14-S3-S2>
- Chan C-H, Yu T-H, Wong K-B (2011) Stabilizing salt-bridge enhances protein thermostability by reducing the heat capacity change of unfolding. *PLoS ONE* 6:e21624
- Chen Y, Lu H, Zhang N et al (2020) PremPS: predicting the impact of missense mutations on protein stability. *PLoS Comput Biol* 16:e1008543. <https://doi.org/10.1371/journal.pcbi.1008543>
- Chorley BN, Campbell MR, Wang X et al (2012) Identification of novel NRF2-regulated genes by CHIP-Seq: influence on retinoid X receptor alpha. *Nucleic Acids Res* 40:7416–7429. <https://doi.org/10.1093/NAR/GKS409>

- Dechend R, Hirano F, Lehmann K et al (1999) The Bcl-3 oncoprotein acts as a bridging factor between NF- κ B/Rel and nuclear co-regulators. *Oncogene* 18(22):3316–3323. <https://doi.org/10.1038/sj.onc.1202717>
- Dehouck Y, Kwasigroch JM, Rooman M, Gilis D (2013) BeAtMuSiC: prediction of changes in protein–protein binding affinity on mutations. *Nucleic Acids Res* 41:W333–W339
- Dunwell JM, Culham A, Carter CE et al (2001) Evolution of functional diversity in the cupin superfamily. *Trends Biochem Sci* 26:740–746. [https://doi.org/10.1016/s0968-0004\(01\)01981-8](https://doi.org/10.1016/s0968-0004(01)01981-8)
- Ghosh S, May MJ, Kopp EB (1998) NF- κ B and Rel proteins: evolutionarily conserved mediators of immune responses. *Annu Rev Immunol* 16:225–260. <https://doi.org/10.1146/annurev.immunol.16.1.225>
- Girdharan S, Srinivasan M (2018) Mechanisms of NF- κ B p65 and strategies for therapeutic manipulation. *J Inflamm Res* 11:407–419. <https://doi.org/10.2147/JIR.S140188>
- Hernandez R, Facelli JC (2021) Understanding protein structural changes for oncogenic missense variants. *Heliyon* 7:e06013. <https://doi.org/10.1016/j.heliyon.2021.e06013>
- Ho BK, Brasseur R (2005) The Ramachandran plots of glycine and pre-proline. *BMC Struct Biol*. <https://doi.org/10.1186/1472-6807-5-14>
- Howe KL, Achuthan P, Allen JJJJ et al (2021) Ensembl 2021. *Nucleic Acids Res* 49:D884–D891. <https://doi.org/10.1093/NAR/GKAA942>
- Hubbard RE, Haider MK (2010) Hydrogen bonds in proteins: role and strength. *Encycl Life Sci*. <https://doi.org/10.1002/9780470015902.A0003011.PUB2>
- Hübner R-H, Schwartz JD, De Bishnu P et al (2009) Coordinate control of expression of Nrf2-modulated genes in the human small airway epithelium is highly responsive to cigarette smoking. *Mol Med* 15:203–219. <https://doi.org/10.2119/molmed.2008.00130>
- Ittisoponpisan S, Islam SA, Khanna T et al (2019) Can predicted protein 3D structures provide reliable insights into whether missense variants are disease associated? *J Mol Biol* 431:2197–2212. <https://doi.org/10.1016/j.jmb.2019.04.009>
- Jamroz M, Orozco M, Kolinski A, Kmiecik S (2013) Consistent view of protein fluctuations from all-atom molecular dynamics and coarse-grained dynamics with knowledge-based force-field. *J Chem Theory Comput* 9:119–125. <https://doi.org/10.1021/ct300854w>
- Jung C, Mock A, Exner J et al (2016) Spatial transcriptome analysis reveals Notch pathway-associated prognostic markers in IDH1 wild-type glioblastoma involving the subventricular zone. *BMC Med* 14:170. <https://doi.org/10.1186/s12916-016-0710-7>
- Kajander T, Kahn PC, Passila SH et al (2000) Buried charged surface in proteins. *Structure* 8:1203–1214. [https://doi.org/10.1016/S0969-2126\(00\)00520-7](https://doi.org/10.1016/S0969-2126(00)00520-7)
- Khanna T, Hanna G, Sternberg MJEE, David A (2021) Missense3D-DB web catalogue: an atom-based analysis and repository of 4M human protein-coding genetic variants. *Hum Genet* 140:805–812. <https://doi.org/10.1007/s00439-020-02246-z>
- Kmiecik S, Gront D, Kolinski M et al (2016) Coarse-grained protein models and their applications. *Chem Rev* 116:7898–7936. <https://doi.org/10.1021/acs.chemrev.6b00163>
- Krieger F, Möglich A, Kieffhaber T (2005) Effect of proline and glycine residues on dynamics and barriers of loop formation in polypeptide chains. *J Am Chem Soc* 127:3346–3352. <https://doi.org/10.1021/JA042798I>
- Kuriata A, Gierut AM, Oleniecki T et al (2018) CABS-flex 2.0: a web server for fast simulations of flexibility of protein structures. *Nucleic Acids Res* 46:W338–W343. <https://doi.org/10.1093/nar/gky356>
- Laimer J, Hofer H, Fritz M et al (2015) MAESTRO—multi agent stability prediction upon point mutations. *BMC Bioinform*. <https://doi.org/10.1186/S12859-015-0548-6>
- Lenardo M, Siebenlist U (1994) Bcl-3-mediated nuclear regulation of the NF- κ B trans-activating factor. *Immunol Today* 15:145–147. [https://doi.org/10.1016/0167-5699\(94\)90308-5](https://doi.org/10.1016/0167-5699(94)90308-5)
- Li Q, Verma IM (2002) NF- κ B regulation in the immune system. *Nat Rev Immunol* 2:725–734. <https://doi.org/10.1038/nri910>
- Licciulli S, Cambiaghi V, Scafetta G et al (2010a) Pirin downregulation is a feature of AML and leads to impairment of terminal myeloid differentiation. *Leukemia* 24:429–437. <https://doi.org/10.1038/LEU.2009.247>
- Licciulli S, Luise C, Zanardi A et al (2010b) Pirin delocalization in melanoma progression identified by high content immuno-detection based approaches. *BMC Cell Biol*. <https://doi.org/10.1186/1471-2121-11-5>
- Lin Y-F, Cheng C-W, Shih C-S et al (2016) MIB: metal ion-binding site prediction and docking server. *J Chem Inf Model* 56:2287–2291. <https://doi.org/10.1021/acs.jcim.6b00407>
- Liu F, Rehmani I, Esaki S et al (2013) Pirin is an iron-dependent redox regulator of NF- κ B. *Proc Natl Acad Sci* 110:9722–9727. <https://doi.org/10.1073/pnas.1221743110>
- Liu T, Zhang L, Joo D, Sun SC (2017) NF- κ B signaling in inflammation. *Signal Transduct Target Ther*. <https://doi.org/10.1038/SIGTRANS.2017.23>
- López-Ferrando V, Gazzo A, de la Cruz X et al (2017) PMut: a web-based tool for the annotation of pathological variants on proteins, 2017 update. *Nucleic Acids Res* 45:W222–W228. <https://doi.org/10.1093/nar/gkx313>
- Lu CH, Lin YS, Chen YC et al (2006) The fragment transformation method to detect the protein structural motifs. *Proteins* 63:636–643. <https://doi.org/10.1002/PROT.20904>
- Machiela MJ, Chanock SJ (2015) LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. *Bioinformatics* 31:3555–3557. <https://doi.org/10.1093/bioinformatics/btv402>
- Maldonado V, Melendez-Zajgla J (2011) Role of Bcl-3 in solid tumors. *Mol Cancer*. <https://doi.org/10.1186/1476-4598-10-152>
- Miyazaki I, Simizu S, Okumura H et al (2010) A small-molecule inhibitor shows that pirin regulates migration of melanoma cells. *Nat Chem Biol* 6:667–673. <https://doi.org/10.1038/nchembio.423>
- Oeckinghaus A, Ghosh S (2009) The NF- κ B family of transcription factors and its regulation. *Cold Spring Harb Perspect Biol*. <https://doi.org/10.1101/CSHPERSPECT.A000034>
- Pahari S, Li G, Murthy AK et al (2020) SAAMBE-3D: predicting effect of mutations on protein–protein interactions. *Int J Mol Sci* 21:2563. <https://doi.org/10.3390/ijms21072563>
- Pang H, Bartlam M, Zeng Q et al (2004) Crystal structure of human pirin: an iron-binding nuclear protein and transcription cofactor. *J Biol Chem* 279:1491–1498. <https://doi.org/10.1074/jbc.M310022200>
- Perez-Dominguez F, Carrillo-Beltrán D, Blanco R et al (2021) Role of pirin, an oxidative stress sensor protein, in epithelial carcinogenesis. *Biology (basel)* 10:1–13. <https://doi.org/10.3390/BIOLOGY10020116>
- Pettersen EF, Goddard TD, Huang CC et al (2021) UCSF ChimeraX: structure visualization for researchers, educators, and developers. *Protein Sci* 30:70–82. <https://doi.org/10.1002/PRO.3943>
- DE Pires V, Ascher DB, Blundell TL (2014) mCSM: predicting the effects of mutations in proteins using graph-based signatures. *Bioinformatics* 30:335–342. <https://doi.org/10.1093/bioinformatics/btt691>
- Ponzoni L, Peñaherrera DA, Oltvai ZN, Bahar I (2020) Rhapsody: predicting the pathogenicity of human missense variants. *Bioinformatics* 36:3084–3092. <https://doi.org/10.1093/BIOINFORMATICS/BTAA127>
- Qi H, Zhang H, Zhao Y et al (2021) MVP predicts the pathogenicity of missense variants by deep learning. *Nat Commun* 12:510. <https://doi.org/10.1038/s41467-020-20847-0>
- Rodrigues CHM, Myung Y, Pires DEV, Ascher DB (2019) mCSM-PP12: predicting the effects of mutations on protein–protein interactions. *Nucleic Acids Res* 47:W338–W344. <https://doi.org/10.1093/NAR/GKZ383>
- Santoro C, Mermod N, Andrews PC, Tjian R (1988) A family of human CCAAT-box-binding proteins active in transcription and DNA replication: cloning and expression of multiple cDNAs. *Nature* 334:218–224. <https://doi.org/10.1038/334218A0>
- Savojardo C, Fariselli P, Martelli PL, Casadio R (2016) INPS-MD: a web server to predict stability of protein variants from sequence and structure. *Bioinformatics* 32:2542–2544. <https://doi.org/10.1093/BIOINFORMATICS/BTW192>
- Schrodinger L, Delano W (2020) PyMOL
- Sherry ST, Ward M-H, Kholodov M et al (2001) dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res* 29:308–311
- Shoily SS, Ahsan T, Fatema K, Sajib AA (2021) Disparities in COVID-19 severities and casualties across ethnic groups around the globe and patterns of ACE2 and PIR variants. *Infect Genet Evol* 92:104888. <https://doi.org/10.1016/j.meegid.2021.104888>
- Sim N-L, Kumar P, Hu J et al (2012) SIFT web server: predicting effects of amino acid substitutions on proteins. *Nucleic Acids Res* 40:W452–W457. <https://doi.org/10.1093/NAR/GKS539>
- Suleman M, Tahirul QM, Saleem S et al (2021) Mutational landscape of pirin and elucidation of the impact of most detrimental missense variants that accelerate the breast cancer pathways: a computational modelling study. *Front Mol Biosci*. <https://doi.org/10.3389/FMOLB.2021.692835>

- Tamura M, Ishizawa M, Isojima T et al (2017) Functional analyses of a novel missense and other mutations of the vitamin D receptor in association with alopecia. *Sci Rep* 7:5102. <https://doi.org/10.1038/s41598-017-05081-x>
- Tina KG, Bhadra R, Srinivasan N (2007) PIC: protein interactions calculator. *Nucleic Acids Res* 35:W473–W476. <https://doi.org/10.1093/NAR/GKM423>
- Wang J, Youkharibache P, Zhang D et al (2020) iCn3D, a web-based 3D viewer for sharing 1D/2D/3D representations of biomolecular structures. *Bioinformatics* 36:131–135. <https://doi.org/10.1093/bioinformatics/btz502>
- Waterhouse A, Bertoni M, Bienert S et al (2018) SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res* 46:W296–W303. <https://doi.org/10.1093/nar/gky427>
- Wendler WMF, Kremmer E, Förster R, Winnacker E-LL (1997) Identification of pirin, a novel highly conserved nuclear protein. *J Biol Chem* 272:8482–8489. <https://doi.org/10.1074/jbc.272.13.8482>
- Yan Y, Tao H, He J, Huang S-Y (2020) The HDock server for integrated protein–protein docking. *Nat Protoc* 15:1829–1852. <https://doi.org/10.1038/s41596-020-0312-x>
- Yoshikawa R, Yanagi H, Hashimoto-Tamaoki T et al (2004) Gene expression in response to anti-tumour intervention by polysaccharide-K (PSK) in colorectal carcinoma cells. *Oncol Rep* 12:1287–1293
- Zhang N, Chen Y, Lu H et al (2020) MutaBind2: predicting the impacts of single and multiple mutations on protein–protein interactions. *iScience* 23:100939. <https://doi.org/10.1016/j.isci.2020.100939>
- Zhang Z, Miteva MA, Wang L, Alexov E (2012) Analyzing effects of naturally occurring missense mutations. *Comput Math Methods Med* 2012:1–15. <https://doi.org/10.1155/2012/805827>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)
