Indian Journal of Biochemistry & Biophysics
Vol. 60, January 2023, pp. 7-25
DOI: 10.56042/ijbb.v60i1.56987

A comparative computational approach on the most deleterious missense variant in Connexin 43 protein and its potent inhibitor analysis

Ramkumar Katturajan[1], Tamma Medha[1], Sakshi Karra[1], Vidya R[2] & Sabina Evan Prince[1]*

[1]Department of Biomedical Sciences, School of Biosciences and Technology; & [2]VIT School of Agri innovations and Advance Learning (VAIAL), VIT Vellore-632 014, Tamil Nadu, India

Received 09 November 2021; *revised 09 December 2022*

Intercellular communication between the cell plays an essential role in cell growth and cell formation, including migration, metabolism, and cell differentiation. Cell function and tissue homeostasis are maintained through gap junction intercellular communication (GJIC), thus regulating connexin hemichannels. Mis regulation of such connexin, especially connexin (Cx) 43, affects a comprehensive process, including cell differentiation, inflammation, and cell death. Mis regulation may be due to the missense variant in Cx43. Thus, we screened the complete set of mutations from public mutational databases and obtained 219 missense variants, which were then classified based on their pathogenicity, functional impact, stability, conservation, and physiochemical properties. Variant L214P was scrutinized to have the most deleterious, which was then modelled using the I-TASSER server and performed molecular docking analysis to screen potent inhibitors. The compound Kanamycin, Ginsenoside, and Astragaloside IV have better interactions with Cx43 mutant with a maximum of 5 hydrogen bonds. Ginsenoside is a compound that follows a Lipinski rule of five. Thus, the result obtained from this study suggests that Ginsenoside would be a better potent inhibitor for native and mutant Cx43.

Keywords: Cx43, L214P, Virtual screening, Variant classification, Molecular docking

Connexins (Cxs) are a multi-gene family of proteins that regulate the intercellular hole intersection of gap junction (GJ) channels to coordinate communication amongst cells[1]. GJ channels are shaped by the docking of two hemichannels, one from each of the two reaching cells. It is presently well established that each hemichannel can work with the nonappearance of docking and subsequently intervening signaling across the plasma film[2]. GJ channels of Hemi channels play an essential role in many aspects of tissue homeostasis within the brain, heart, and other tissues, as evidenced by the link between a growing list of human illnesses and changes in connexin characteristics[3]. It is fundamental for some physiological cycles, like coordinated depolarization of cardiac muscle, proper embryonic development, and the conducted response in the microvasculature.

Consequently, transformations in connexin-encoding qualities can prompt functional and formative anomalies. Twenty-one different Cxs have been identified and studied, which consist of four transmembrane helices (TM1-TM4), two extracellular loops (ECL1 and ECL2), an N-terminal helix, and a large carboxy-terminal domain[4]. Among all the Cxs, Cx43 is widely distributed in almost all the cell types in most organs and is significantly expressed under disease conditions[5].

Under stress situations, the activity of the hemi channels changes that are entailing moving molecules such as $Ca2^{+}$, ATP, NAD^{+}, and glutamate to another cell and inducing numerous physiological responses[5]. Mutations in 10 different human Cxs have been related to 28 hereditary disorders. Cx43 mutations were responsible for more than six illnesses[3]. As a result, many mechanistic studies have been conducted on this Cx43. Concerning the importance of Cxs mutations, connexinopathies have been identified, termed diseases related to the Cxs mutations. Mutations in Cx43 were reported to have several genetic disorders, including oculodentodigital dysplasia, palmoplantar keratoderma, congenital alopecia, hyperkeratosis, leukonychia, erythron-keratodermia variabilis et progressive and linear verrucous epidermal nevus[6-8]. In addition, Cx43 mutations are also associated with several heart diseases where the principal role of Cx43 in the myocardium is to regulate the rapid and coordinated excitation-contraction coupling mechanisms[9].

*Correspondence:
Phone: 91-416-2202324; +91-9080494445 (Mob)
E-mail: eps674@gmail.com; epsabina@vit.ac.in

Moreover, several reports have suggested that Cx43 mutants lack a C-terminal tail which results in inhibited cell division and failure to form GJ shown to have down regulated cell growth[10]. Nonetheless, there is literature to assist in identifying the most harmful or significant mutations responsible for disease causation and development[11]. Rangasamy *et al*., 2021 has recently reported that computational approach to predict mutations are essential in scrutinizing the most significant disease-causing mutation[12].

The computational design approach comprises virtual screening and molecular docking that has manifested trustable evidence in drug developments and definite outcomes[13,14]. Cx43 has been of foremost importance in various disease conditions, and the significance of the related mutations has been analyzed using various web-based tools. The present study examines the first analysis of the mutational landscape of Cx43 in association with a virtual screening of Cxs inhibitors on most pathogenic mutation L214P. Primarily, we screened the complete set of mutations from mutational databases and classified them based on their pathogenicity, functional impact, stability, conservation, and physiochemical properties. A native and mutant model with the desired variation was modeled using the I-TASSER server, and the Cx inhibitor was screened based on the literature survey (Table 4). Molecular docking was performed to find the potent desired Cx43 mutant inhibitor. This computational strategy for discovering harmful mutations and screening for effective inhibitors of those mutations may soon contribute to the creation of tailored medicine.

Materials and Methods

Collection of data

The mutations and their combined information for the Cx43 Missense variant were retrieved from databases like COSMIC (https://cancer.sanger.ac.uk/cosmic), HGMD (http://www.hgmd.cf.ac.uk/ac/index.php), and literature survey. Sequence information of Cx43 in FASTA format was retrieved from Uniport KB (https://www.uniprot.org/).

Pathogenicity analysis of missense variant

Calculation methods were used to understand the impact of variations on proteins which is vital for classifying and prioritizing pathogenic in neutral single-nucleotidevariations[15].

Meta-SNP (http://snps.biofold.org/meta-snp/) is a web-based server for many genome-related studies, which improves the ability to detect more heterogeneity of associations and investigate the consistency from different data sets and research populations. It integrates the best performance prediction algorithms to classify the pathogenicity of protein variants. In addition, the algorithm integrates with various other algorithms, including SNAP prediction, PhD SNP prediction, PANTHER prediction, and SIFT prediction. The predictor outputs the probability that the specified variant is associated with the disease. Here a score of > 0.5 establishes that a particular mutation induces the disease.

Functional impact analysis of the missense variant was done with the help of the Mutation Assessor server (http://mutationassessor.org/r3/). It is a network-dependent application that leverages disease-related Online Mendelian Inheritance in Man (OMIM) and polymorphism information to assess the effects of changes in a single-point amino acid change. The mutation assessor uses the Uniport protein sequence to generate its Multiple sequence alignment (MSA). It then splits based on the boundaries of the Uniport and Pfam domains to generate a 3D structure using a sorted product set and subfamily set. The segmented MSA was created to identify evolutionarily conserved locations that contribute to the specificity of protein function. Conservation scores are combined with specificity assessments to determine functional impact. As an outcome, mutants classified as "neutral" or "low" are not expected to affect protein function, whereas mutants classified as "medium" or "high" are functional and are expected to bring about changes.

Structure stability analysis

Structure stability analysis helps determine whether a protein will be in a native folded conformation or a denatured state. It refers to physical stability (thermodynamic) and not chemical stability. Mutations in the protein frequently change the stability of the protein[16]. Here, the difference in free energy ($\Delta\Delta G$) between the mutant (ΔGm) and the wild-type protein (ΔGw) is a measure of how a particular mutation affects the stability ofthe protein. A positive $\Delta\Delta G$ value shows a stabilizing mutation. We used different computational methods like DUET, MUpro, INPS-MD, i-Mutant 2.0, and Dyna Mut.

DUET (http://biosig.unimelb.edu.au/duet/stability) is a web server for integrated computer access to study missense mutations in proteins. It combines two

complementary approaches (mCSM and SDM) of consensus prediction obtained by blending results of particular methods in prediction optimized using SVM (Support Vector Machines). DUET improves the overall accuracy of the forecast compared to either technique by itself. By selectively combining the two methods, it far surpasses another integrated approach that combines the seven methods.

Mupro (http://mupro.proteomics.ics.uci.edu/) is a set of machine learning programs for predicting the effects of single-site amino acid mutations on protein stability. Two machine learning methods were developed, which are SVM and Neural Networks. An advantage of the method is that it does not require a tertiary structure to predict changes in protein stability.

INPS-MD (https://inpsmd.biocomp.unibo.it/inpsSuite/default/index3D) (Impact of Nonsynonymous mutations on Protein Stability Multi Dimension) is a web server designed to predict changes in protein stability duringa single point mutation. Currently, two versions of predictive variables are used. INPS prediction variable from the sequence: Prediction of the impact of nonsynonymous Single Nucleotide Polymorphisms (nsSNPs) on protein stability on protein stability sequence. INPS Predictor of protein 3D structure: Predicting the impact of asynchronous ns SNP for protein stability, starting with protein Structure.

I-Mutant2.0 (https://folding.biofold.org/cgi-bin/i-mutant2.0.cgi) is an SVM-based tool for automatically predicting changes in protein stability due to single-point mutations. These predictions are performed starting from the structure of the protein or, more importantly, the protein's sequence. IMutant2.0 is a classifier that predicts signs of changes in the stability of a protein upon mutation, which can be used as a regression estimator to predict the relevant $\Delta\Delta G$ values. The web server passes the protein array into its raw format.

Dyna Mut(http://biosig.unimelb.edu.au/dynamut/prediction) implements two well-established, normal-mode approaches to web servers that sample structures, analyze and visualize protein dynamics, and determine protein dynamics and stability due to vibrational entropy changes. It accommodates graph-based signatures with normal mode dynamics to achieve consensus predictions about the effects of mutations on protein stability. It also offers a comprehensive suite for protein motility, flexibility analysis, and visualization via a free, user-friendly web server[17].

Predicted binding site

COACH-D (https://yanglab.nankai.edu.cn/COACH-D/) is a method for accessing the meta server for predicting protein-ligand binding sites. It starts with a specific target protein structure and uses two comparative methods, TMSITE and SSITE, to generate a prediction of the complementary ligand-binding site. This method recognizes ligand binding templates in the functional database of BioLiP proteins by comparing binding-specific sub structures and sequence profiles. Initially, five separate ways are used to predict the ligand-binding pockets and residues. The template is then docked in the binding pocket. One of the significant improvements of COACH-D over COACH is that it uses Auto Dock Vina, an efficient molecular docking algorithm, to improve the ligand binding pose and make it physically and more realistic. The major conclusion is:Predicted 3D structural model submitted with a protein sequence, Top 5 protein-ligand binding pockets and binding residues in each pocket, top 5 protein-ligand complexes, submitted ligand docking structures, ligand docking from template structures, top 5 protein-ligand complex structures[18].

Modelling native and variant protein

I-TASSER (https://zhanglab.dcmb.med.umich.edu/I-TASSER/) (Iterative Threading Assembly Refinement) is a progressive way to deal with protein structure forecast and structure-based function annotation. First, structural templates in the PDB by a multi-threaded approach were identified using LOMETS full-length atomic models built by iterative template-based fragment assembly simulations. Next, the 3D model is re-threaded through the protein function database BioLiP to derive insight into the target function. It was recently ranked as the No.1 server for protein structure prediction in CASP7, CASP8, CASP9, CASP10, CASP11, CASP12, CASP13, and CASP14 experiments throughout society. It has also been evaluated with CASP9 for functional prediction. The server is under dynamic improvement, fully intent on giving the most exact protein design and capacity forecasts utilizing modern algorithms.

Loop refined was done using the HADDOCK server (https://wenmr.science.uu.nl/haddock2.4/refinement/1). High Ambiguity Driven Protein Docking) is an information-driven, flexible docking approach for the modelling of biomolecular complexes. HADDOCK differs from the abiniteo docking method in that it

encodes information at a protein interface identified or predicted in ambiguous interaction inhibition (AIR) to drive the docking process. It is also possible to define specific, clear distance limits(suchas MS cross-links),NMR residual dipole coupled pseudo-contact shifts, frozen EM maps, and many other experimental data supports. HADDOCK Proteins can deal with a massive class of displaying issues, including proteins, protein-nucleic acids, protein-ligand buildings, and multi-body (n>2) gatherings. HADDOCK is one of the flagship software for biomolecular research at the EU H2020 Bio excel center of excellence[19].

Structures were validated by the Ramachandran plot server (https://zlab.umassmed.edu/bu/rama/index.pl). This server displays Ramachandran plots against the background of whiplash probabilities, and the method server display color Ramachandran Plot. According to DSSP, blue means helix, red means strand, and green means turn-and-loop. The plotline shows the priority area. The outline surrounds the area where 90% of crosses of the same color are found. Lines inside show 50% area.

Conservative sequence analysis

ConSurf server (https://consurf.tau.ac.il/) is a bioinformatics tool for estimating the evolutionary storage of amino/nucleic acid positions of protein/DNA/RNA molecules according to phylogenetic relationships between homologous sequences. The extent to which the position of an amino acid (or nucleic acid) is evolutionarily conserved (*i.e.*, its rate of evolution) is highly dependent on the structural and functional significance. Therefore, analysis of position storage between members of the same family often clarifies the importance of each position to the structure or function of a protein (or nucleic acid). ConSurf estimates evolutionary rates by considering the similarities between amino acids (nucleic acids) that are reflected in alternative matrices according to the evolutionary relevance between proteins (DNA/RNA) and their homologs. One of the upsides of ConSurf over different techniques is that it precisely computes the pace of development utilizing either the exact Bayes strategy or the most extreme probability (ML) strategy[20].

Preparation of ligands

Thirty-six compounds were scrutinized by a literature survey based on the inhibitory effect on Cx43[4,5]. The information and SDF format of the 3D structure of the compounds were obtained from the PubChem database (https://pubchem.ncbi.nlm.nih.gov/). SDF formatted compounds were further converted to PDBQT format by Open Babel software which was used for docking[21].

Molecular docking

Molecular docking was performed by using the Autodock Vina software. Water from the native and L214P mutated proteins was removed, and polar hydrogen, solvation, and charges were added to the proteins. Affinity maps with grid points were fixed for the active binding sites of the proteins by using the Auto Grid program. A Lamarckian genetic algorithm was used to perform protein-ligand docking in Autodock vina. The results obtained from 10 different runs for each docking complex, among the highest binding energy complexes, were visualized by Pymol and Discovery studio software.

Results

Metadata and disease-causing missense

A list of 249 missense variants for Cx43 was retrieved from public databases and literature review, followed by missense repetition removed and finalized to 219 missense variants. These missenses were then screened for pathogenicity analysis using a meta-SNP web-based server which includes PANTHER, PhD-SNP, SIFT, SNAP, and meta-SNP server (Fig. 1). As a result, among 219 missense, 52 missense variants were found to have deleterious in all the servers, which were taken to functional impact analysis (Table 1).

Function impact analysis of selected Cx43 mutations

The functional impact of the selected 52 missenses was examined using a mutation assessor server. As a result, 24 mutations were predicted to have a significant impact, and 22 missenses were shown to

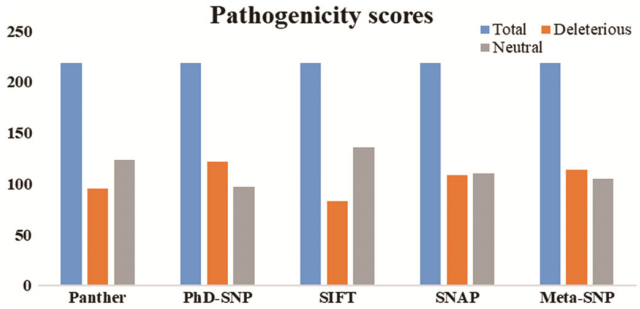


Fig. 1 — Deleterious and neutral mutation screening of Cx43 missense

Table 1 — Cx43 mutations were classified as deleterious or neutral using the meta-SNP server

		Panther		PhD-SNP		SIFT		SNAP		Meta-SNP	
Sl.no	Mutations (AA)	Disease	score	Disease	score	Disease	score	Disease	score	Disease	score
1	P363L	YES	0.603	NO	0.233	NO	0.37	NO	0.465	NO	0.344
2	S369N	NO	0.437	NO	0.123	NO	0.56	NO	0.29	NO	0.145
3	R370S	NO	0.294	NO	0.12	NO	0.74	NO	0.315	NO	0.104
4	R370H	NO	0.229	NO	0.087	NO	0.12	NO	0.45	NO	0.07
5	E381K	NO	0.239	NO	0.047	NO	0.85	NO	0.19	NO	0.057
6	W4C	YES	0.829	YES	0.66	YES	0	YES	0.765	YES	0.747
7	A6T	NO	0.364	NO	0.079	NO	0.07	NO	0.145	NO	0.275
8	K13N	NO	0.459	YES	0.5	NO	0.27	NO	0.225	NO	0.446
9	Y17C	YES	0.774	YES	0.855	YES	0	YES	0.725	YES	0.811
10	G21E	NO	0.29	YES	0.683	YES	0	YES	0.645	YES	0.749
11	W25R	YES	0.639	YES	0.978	YES	0	YES	0.795	YES	0.91
12	R33Q	YES	0.997	YES	0.935	YES	0	YES	0.815	YES	0.952
13	L37P	YES	0.716	YES	0.962	YES	0	YES	0.61	YES	0.817
14	S43T	NO	0.318	NO	0.275	NO	0.25	NO	0.235	NO	0.458
15	S43L	NO	0.469	YES	0.626	NO	0.09	NO	0.405	YES	0.516
16	Q49K	NO	0.262	YES	0.721	YES	0.02	YES	0.665	YES	0.561
17	R53G	NO	0.277	NO	0.323	NO	0.06	NO	0.48	NO	0.469
18	R53C	YES	0.556	NO	0.494	YES	0.01	YES	0.56	YES	0.696
19	R53H	NO	0.284	NO	0.169	YES	0.05	NO	0.4	NO	0.302
20	R53L	NO	0.207	NO	0.337	NO	0.1	NO	0.425	NO	0.459
21	G60C	YES	0.999	YES	0.945	YES	0	YES	0.705	YES	0.921
22	C61S	NO	0.468	YES	0.946	YES	0.03	YES	0.64	YES	0.785
23	K68N	NO	0.405	NO	0.323	NO	0.59	NO	0.32	NO	0.466
24	P71T	YES	0.998	YES	0.952	YES	0	YES	0.755	YES	0.896
25	R76C	YES	0.812	YES	0.942	YES	0	YES	0.835	YES	0.924
26	V79F	YES	0.524	YES	0.9	YES	0.02	YES	0.59	YES	0.619
27	F84C	YES	0.783	YES	0.931	YES	0.04	YES	0.62	YES	0.749
28	V85G	YES	0.576	YES	0.952	YES	0	YES	0.745	YES	0.898
29	P88L	YES	0.641	YES	0.97	YES	0	YES	0.735	YES	0.877
30	A94D	YES	0.552	YES	0.851	YES	0.01	YES	0.6	YES	0.814
31	A94V	NO	0.408	YES	0.532	NO	0.26	NO	0.27	NO	0.455
32	Y98S	YES	0.576	YES	0.894	YES	0	YES	0.625	YES	0.77
33	R101Q	NO	0.478	YES	0.668	NO	0.13	NO	0.49	YES	0.529
34	E103K	NO	0.362	YES	0.706	NO	0.11	YES	0.54	YES	0.584
35	E104D	NO	0.276	YES	0.562	YES	0.03	NO	0.475	YES	0.52
36	E110K	NO	0.417	NO	0.219	NO	0.42	NO	0.49	NO	0.411
37	E112K	NO	0.417	NO	0.276	NO	0.69	NO	0.325	NO	0.443
38	K114N	NO	0.439	NO	0.212	NO	0.49	NO	0.34	NO	0.37
39	A116V	NO	0.054	NO	0.136	NO	0.28	NO	0.415	NO	0.302
40	G120A	NO	0.423	NO	0.39	NO	0.5	NO	0.46	NO	0.45
41	V123A	NO	0.34	NO	0.141	NO	0.55	NO	0.38	NO	0.168
42	M125I	NO	0.126	NO	0.16	NO	0.13	NO	0.385	NO	0.228
43	L127W	YES	0.831	YES	0.505	NO	0.13	NO	0.405	YES	0..552
44	L127F	NO	0.45	NO	0.254	NO	0.48	NO	0.205	NO	0.429
45	Q129E	NO	0.159	NO	0.19	NO	1	NO	0.185	NO	0.301
46	K133N	NO	0.431	NO	0.39	NO	0.37	NO	0.205	NO	0.463
47	G138C	YES	0.805	NO	0.418	NO	0.05	NO	0.48	YES	0.625
48	E141K	NO	0.411	YES	0.545	NO	0.76	NO	0.28	NO	0.459
49	V145L	NO	0.223	NO	0.361	NO	0.35	NO	0.175	NO	0.433
50	G149E	YES	0.996	YES	0.872	NO	0.12	YES	0.59	YES	0.866

(*Contd.*)

Table 1 — Cx43 mutations were classified as deleterious or neutral using the meta-SNP server (*Contd.*)

		Panther		PhD-SNP		SIFT		SNAP		Meta-SNP	
Sl.no	Mutations (AA)	Disease	score	Disease	score	Disease	score	Disease	score	Disease	score
101	S306N	NO	0.126	NO	0.317	NO	0.31	NO	0.32	NO	0.3
102	S314R	NO	0.494	NO	0.387	NO	0.41	NO	0.325	NO	0.291
103	Q317R	NO	0.324	NO	0.299	NO	0.5	NO	0.375	NO	0.293
104	R319Q	NO	0.423	NO	0.332	NO	0.29	NO	0.475	NO	0.315
105	A323V	NO	0.302	NO	0.195	NO	0.3	NO	0.46	NO	0.244
106	G324E	YES	0.583	YES	0.58	NO	1	NO	0.41	NO	0.448
107	S328P	YES	0.525	YES	0.62	NO	0.29	NO	0.2	NO	0.445
108	S330F	YES	0.594	YES	0.572	NO	0.7	NO	0.34	NO	0.462
109	A332T	NO	0.347	NO	0.269	NO	0.58	NO	0.325	NO	0.265
110	D336V	YES	0.633	NO	0.368	NO	0.28	NO	0.48	NO	0.45
111	D340N	NO	0.225	NO	0.456	NO	0.35	NO	0.26	NO	0.342
112	D340Y	YES	0.619	YES	0.673	NO	1	NO	0.25	YES	0.532
113	Q342R	NO	0.107	NO	0.318	NO	0.33	YES	0.505	NO	0.294
114	N343H	NO	0.374	NO	0.19	NO	0.15	NO	0.435	NO	0.246
115	K345R	NO	0.242	NO	0.108	NO	0.49	NO	0.215	NO	0.088
116	L347P	YES	0.642	YES	0.653	NO	0.27	NO	0.435	NO	0.455
117	A348T	NO	0.092	NO	0.127	NO	0.59	NO	0.23	NO	0.11
118	G350E	YES	0.574	NO	0.44	NO	1	NO	0.315	NO	0.458
119	L356R	YES	0.623	YES	0.712	NO	0.52	NO	0.3	YES	0.539
120	V359L	NO	0.111	NO	0.073	NO	0.74	NO	0.32	NO	0.065
121	D360E	NO	0.274	NO	0.302	NO	1	NO	0.225	NO	0.256
122	T19A	NO	0.142	YES	0.853	YES	0	YES	0.77	YES	0.786
123	V28I	NO	0.166	NO	0.498	NO	0.24	YES	0.51	NO	0.429
124	A44V	NO	0.367	NO	0.08	NO	1	NO	0.215	NO	0.266
125	E227D	YES	0.993	YES	0.883	YES	0.04	YES	0.65	YES	0.576
126	P283L	YES	0.637	YES	0.564	NO	0.05	YES	0.645	YES	0.51
127	T290D	YES	0.515	NO	0.429	NO	0.66	YES	0.56	NO	0.372
128	Y17S	YES	0.531	YES	0.792	YES	0	YES	0.66	YES	0.76
129	G143S	NO	0.455	YES	0.626	NO	0.58	NO	0.405	NO	0.477
130	G138I	YES	0.785	NO	0.294	NO	0.2	NO	0.37	NO	0.464
131	S18P	YES	0.994	YES	0.913	YES	0	YES	0.805	YES	0.92
132	G21R	NO	0.351	YES	0.741	YES	0	YES	0.655	YES	0.758
133	G22E	YES	0.565	YES	0.962	YES	0	YES	0.845	YES	0.933
134	G60S	YES	0.997	YES	0.913	YES	0	YES	0.685	YES	0.879
135	G138R	YES	0.716	NO	0.435	NO	0.53	NO	0.435	YES	0.566
136	V96M	NO	0.451	YES	0.73	YES	0.03	YES	0.645	YES	0.682
137	I31M	NO	0.407	YES	0.814	YES	0	YES	0.66	YES	0.681
138	G8V	NO	0.348	YES	0.628	YES	0.04	YES	0.645	NO	0.454
139	G2V	YES	0.59	NO	0.447	YES	0	YES	0.62	YES	0.597
140	D3N	NO	0.257	NO	0.113	NO	0.15	NO	0.25	NO	0.231
141	S5C	YES	0.648	NO	0.335	YES	0	NO	0.465	YES	0.596
142	L7V	NO	0.3	NO	0.483	YES	0.01	YES	0.77	YES	0.578
143	L11I	NO	0.274	YES	0.518	YES	0.04	YES	0.645	NO	0.454
144	L11P	YES	0.743	YES	0.907	YES	0	YES	0.775	YES	0.818
145	L11F	NO	0.426	YES	0.734	YES	0.01	YES	0.675	YES	0.648
146	G22R	YES	0.652	YES	0.962	YES	0.02	YES	0.84	YES	0.93
147	K23T	NO	0.437	YES	0.893	YES	0	YES	0.735	YES	0.807
148	V24A	NO	0.327	YES	0.765	YES	0	YES	0.64	YES	0.706
149	W25C	YES	0.806	YES	0.974	YES	0	YES	0.775	YES	0.892
150	S27P	NO	0.442	YES	0.955	YES	0.04	YES	0.595	YES	0.623

(*Contd.*)

Table 1 — Cx43 mutations were classified as deleterious or neutral using the meta-SNP server (*Contd.*)

		Panther		PhD-SNP		SIFT		SNAP		Meta-SNP	
Sl.no	Mutations (AA)	Disease	score	Disease	score	Disease	score	Disease	score	Disease	score
151	A40V	NO	0.367	YES	0.637	NO	0.09	YES	0.585	YES	0.501
152	V41L	NO	0.177	NO	0.437	YES	0.02	YES	0.51	NO	0.497
153	E42L	YES	0.527	YES	0.793	YES	0.01	YES	0.66	YES	0.709
154	E42Q	NO	0.406	YES	0.687	YES	0.03	YES	0.545	YES	0.528
155	D47H	YES	0.998	YES	0.878	YES	0	YES	0.75	YES	0.889
156	E48L	YES	0.527	YES	0.818	YES	0	YES	0.715	YES	0.744
157	Q49L	NO	0.421	YES	0.781	YES	0	YES	0.655	YES	0.661
158	Q49P	YES	0.56	YES	0.88	YES	0	YES	0.655	YES	0.761
159	Q49E	NO	0.299	YES	0.666	YES	0.02	YES	0.66	YES	0.605
160	N55D	NO	0.344	YES	0.615	NO	0.15	YES	0.62	NO	0.441
161	Q58H	YES	0.588	YES	0.845	YES	0	YES	0.71	YES	0.779
162	P59H	YES	0.999	YES	0.893	YES	0	YES	0.78	YES	0.937
163	P59A	YES	0.997	YES	0.796	YES	0.02	YES	0.7	YES	0.851
164	S69Y	YES	0.626	YES	0.706	NO	1	NO	0.28	YES	0.578
165	H74L	YES	0.522	YES	0.908	NO	0.13	YES	0.615	YES	0.598
166	R76S	NO	0.491	YES	0.926	YES	0	YES	0.78	YES	0.817
167	R76H	YES	0.621	YES	0.932	YES	0.02	YES	0.82	YES	0.848
168	V85M	YES	0.506	YES	0.882	YES	0.01	YES	0.73	YES	0.779
169	S86Y	YES	0.683	YES	0.954	YES	0	YES	0.725	YES	0.886
170	L90V	NO	0.304	YES	0.718	NO	0.13	NO	0.485	YES	0.55
171	H95R	NO	0.468	YES	0.909	YES	0	YES	0.79	YES	0.836
172	V96E	YES	0.571	YES	0.921	YES	0.01	YES	0.8	YES	0.856
173	V96A	NO	0.312	NO	0.322	NO	0.83	NO	0.42	NO	0.45
174	Y98C	YES	0.802	YES	0.904	YES	0	YES	0.67	YES	0.791
175	R101L	YES	0.564	YES	0.831	NO	0.07	NO	0.47	YES	0.704
176	K102N	NO	0.439	NO	0.42	NO	0.36	NO	0.47	NO	0.475
177	L106P	YES	0.768	YES	0.712	NO	0.32	NO	0.46	YES	0.562
178	L106R	YES	0.683	NO	0.383	NO	1	NO	0.27	YES	0.576
179	E110D	NO	0.32	NO	0.206	NO	0.17	YES	0.525	NO	0.387
180	L113P	YES	0.768	YES	0.638	NO	0.33	YES	0.62	YES	0.596
181	I130T	NO	0.432	NO	0.402	NO	0.56	NO	0.475	NO	0.439
182	K134N	NO	0.431	NO	0.477	NO	0.51	NO	0.415	NO	0.474
183	K134E	NO	0.347	NO	0.496	NO	1	NO	0.37	NO	0.46
184	G138S	YES	0.517	NO	0.128	NO	0.73	NO	0.23	NO	0.159
185	G138D	YES	0.643	NO	0.478	NO	0.62	NO	0.49	YES	0.546
187	G143D	YES	0.582	YES	0.875	NO	0.48	YES	0.66	YES	0.714
188	K144E	NO	0.347	YES	0.813	NO	0.05	YES	0.58	YES	0.6
189	V145G	YES	0.545	YES	0.651	NO	0.38	NO	0.275	NO	0.472
190	M147T	NO	0.473	YES	0.5	NO	0.23	YES	0.55	YES	0.643
191	R148Q	NO	0.222	YES	0.608	NO	0.23	YES	0.615	NO	0.478
192	R148G	NO	0.413	YES	0.824	NO	0.14	YES	0.72	YES	0.765
193	R153Q	NO	0.416	YES	0.651	NO	0.32	NO	0.465	YES	0.513
194	T154N	YES	0.526	YES	0.898	YES	0.02	YES	0.715	YES	0.798
195	T154A	NO	0.292	YES	0.687	NO	0.14	YES	0.51	YES	0.531
196	P193L	YES	0.686	YES	0.875	YES	0.04	YES	0.635	YES	0.637
197	H194P	YES	0.657	YES	0.868	YES	0.01	YES	0.705	YES	0.806
198	S201Y	YES	0.719	YES	0.914	YES	0	YES	0.78	YES	0.829
199	S201F	YES	0.701	YES	0.899	YES	0	YES	0.785	YES	0.831
200	R202H	YES	0.66	YES	0.862	YES	0.01	YES	0.755	YES	0.672
201	K206R	YES	0.992	YES	0.879	YES	0	YES	0.69	YES	0.858

(*Contd.*)

Table 1 — Cx43 mutations were classified as deleterious or neutral using the meta-SNP server (*Contd.*)

		Panther		PhD-SNP		SIFT		SNAP		Meta-SNP	
Sl.no	Mutations (AA)	Disease	score	Disease	score	Disease	score	Disease	score	Disease	score
202	V216L	NO	0.272	YES	0.778	YES	0.03	YES	0.66	YES	0.605
203	S220Y	YES	0.719	YES	0.896	YES	0.01	YES	0.685	YES	0.779
204	R239Q	NO	0.317	NO	0.457	NO	0.36	YES	0.58	NO	0.451
205	R239W	YES	0.809	YES	0.701	NO	0.18	YES	0.6	YES	0.733
206	S251T	NO	0.122	NO	0.084	NO	0.6	NO	0.38	NO	0.132
207	A253P	NO	0.069	NO	0.389	NO	0.28	NO	0.36	NO	0.348
208	A253V	NO	0.161	NO	0.207	NO	0.34	YES	0.515	NO	0.272
209	G261W	YES	0.745	YES	0.661	YES	0.02	YES	0.64	YES	0.711
210	S272P	YES	0.545	YES	0.72	NO	0.5	NO	0.25	YES	0.503
211	A276P	YES	0.568	NO	0.482	NO	0.29	NO	0.265	NO	0.484
212	T290N	NO	0.497	NO	0.391	NO	0.74	NO	0.335	NO	0.311
213	A323G	NO	0.274	NO	0.14	NO	0.4	NO	0.395	NO	0.13
214	T326I	NO	0.429	NO	0.278	NO	0.26	NO	0.22	NO	0.338
215	E352G	NO	0.441	NO	0.495	NO	0.3	YES	0.61	YES	0.682
216	R362Q	NO	0.416	YES	0.567	NO	0.59	YES	0.565	YES	0.636
217	S364P	NO	0.213	NO	0.232	NO	0.24	NO	0.41	NO	0.264
218	S365N	NO	0.437	NO	0.377	NO	0.33	NO	0.36	NO	0.369
219	R376Q	NO	0.416	YES	0.507	NO	0.51	YES	0.595	YES	0.611

Table 2 — Functional impact of selected missense in Cx43 protein

Sl.no	Mutations	FI score	VC score	VS score	Functional impact	Sl.no	Mutations	FI score	VC score	VS score	Functional impact
1	W4C	3.81	5.14	2.48	high	27	E205K	4.04	5.68	2.4	high
2	L11F	3.21	3.94	2.48	medium	28	D259Y	1.245	1.39	1.1	low
3	Y17C	2.35	4.01	0.69	medium	29	E227D	2.63	2.86	2.4	medium
4	G22E	3.83	5.18	2.48	high	30	Y17S	1.655	2.62	0.69	low
5	W25R	3.815	5.15	2.48	high	31	S18P	3.805	5.13	2.48	high
6	R33Q	3.84	5.2	2.48	high	32	G60S	3.84	5.2	2.48	high
7	L37P	3.51	4.62	2.4	high	33	L11P	3.76	5.04	2.48	high
8	P71T	3.435	4.47	2.4	medium	34	G22R	3.83	5.18	2.48	high
9	G60C	3.84	5.2	2.48	high	35	W25C	3.47	4.46	2.48	medium
10	R76C	3.785	5.17	2.4	high	36	E42L	Neutral	Neutral	Neutral	Neutral
11	R76H	2.885	4.07	1.7	medium	37	D47H	3.48	4.48	2.48	medium
12	V79F	1.955	2.81	1.1	medium	38	E48L	Neutral	Neutral	Neutral	Neutral
13	F84C	2.725	3.84	1.61	medium	39	Q49P	3.65	5	2.3	high
14	V85M	3.61	5.02	2.2	high	40	Q58H	3.64	5.08	2.2	high
15	V85G	3.61	5.02	2.2	high	41	P59H	3.47	4.46	2.48	medium
16	P88L	3.835	5.19	2.48	high	42	P59A	3.815	5.15	2.48	high
17	A94D	1.495	2.3	0.69	low	43	S86Y	3.38	4.68	2.08	medium
18	Y98S	2.51	4.33	0.69	medium	44	V96E	3.48	4.66	2.3	medium
19	T154N	3.005	3.71	2.3	medium	45	Y98C	2.51	4.33	0.69	medium
20	Y177C	3.95	5.5	2.4	high	46	P193L	2.805	3.21	2.4	medium
21	G178E	3.97	5.54	2.4	high	47	H194P	3.385	4.47	2.3	medium
22	H194L	2.835	3.37	2.3	medium	48	S201Y	3.985	5.57	2.4	high
23	R202H	3.215	4.03	2.4	medium	49	S201F	3.985	5.57	2.4	high
24	T204K	2.925	3.45	2.4	medium	50	K206R	4.045	5.69	2.4	high
25	T204M	2.815	3.23	2.4	medium	51	S220Y	3.67	4.94	2.4	high
26	L214P	3.405	4.51	2.3	medium	52	G261W	1.59	1.79	1.39	low

have a medium impact on protein structure and functionality, which was calculated based on FI, VC, and VS scores. Furthermore, four mutations were found to have a low impact,and two mutations were predicted to have no impact on Cx43 protein functionality, respectively (Fig. 2 and Table 2). Thus, among 52 mutations,the functional impacts, including high and medium of 46 mutations, were taken to further analysis.

Stability analysis of Selected Cx43 mutation

From the mutation functional impact analysis, 46 mutations in the Cx43 protein were predicted to have a high and medium impact on protein functions which

was selected for stability analysis. The selected Cx43 mutant stability was analyzed using ΔΔG analysis-based servers such as DUET, Mupro, INPS-MD, I-Mutant2.0, and Dyna Mut (Table 3). As a result, the stability analysis servers, including mCSM, SDM, DUET, Mupro, INPS-MD, I MUTANT, I Stabilizing, ENCOM, and Dyna MUT predicted destabilizing mutations as 40, 30, 38, 42, 25, 39, 37, 25 and 22 respectively (Fig. 3). From the stability prediction, eight mutations (R76H, V79F, F84C, V85G, Y177C, L214P, G60S, and L11P) are commonly destabilizing in individual servers.

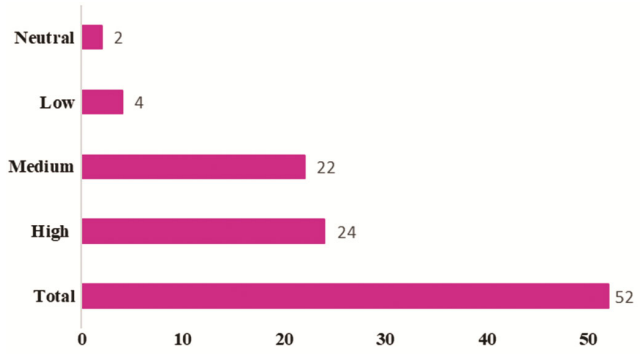


Fig. 2 — Functional impact prediction of Cx43 protein mutations

Binding pocket prediction

Ligand binding site prediction is important for protein regulation. Thus, modelled Cx43 native protein was subjected to predict binding pockets using the COACH-D server. COACH-D result analysis revealed native Cx43 shown to bind with ligands (FE, ZN, 0F1 and PTY) via 22 residues namely, ARG33, LEU37, VAL41, CYS54, CYS61, HIS74, ILE82, VAL166, PHE169, LEU170, GLN173, CYS187, CYS192, CYS198, ILE210, MET213, LEU214, SER217, LEU218, SER220, LEU221 and ALA222 which was considered for binding sites in native and mutant Cx43.

Of the selected eight mutations from stability analysis, a mutation cc was screened to have played a part in binding pockets. Thus, a mutation L214P was then selected for conservative structural analysis.

Table 3 — Stability analysis of Cx43 selected missenses

				DUET				Mupro		INPS–MD		I MUTANT 2.0 SEQ				Dyna Mut			
		mCSM		SDM		DUET		Mupro		INPS–MD		I MUTANT 2.0 SEQ		I Stabilizing		ENCOM		Dyna Mut	
Sl. no	Mutations	ΔΔG	Stability	ΔΔG	Stability	ΔΔG	Stability	ΔΔG	Stability	ΔΔG	Stability	ΔΔG	Stability	ΔΔG	Stability	ΔΔG	Stability	ΔΔG	Stability
1	W4C	0.254	Stabilizing	–0.15	Destabilizing	0.39	Stabilizing	–0.153	Destabilizing	–1.56725	Destabilizing	–1.55	Destabilizing	0.643	Stabilizing	0.002	Stabilizing	0.447	Stabilizing
2	L11F	–0.77	Destabilizing	–0.69	Destabilizing	–0.844	Destabilizing	–1.213	Destabilizing	–1.1666655	Destabilizing	–1	Destabilizing	0.91	Destabilizing	0.379	Stabilizing	0.556	Stabilizing
3	Y17C	–0.46	Destabilizing	–0.07	Destabilizing	–0.29	Destabilizing	–0.556	Destabilizing	–1.496945	Destabilizing	–0.82	Stabilizing	0.55	Destabilizing	–0.699	Destabilizing	–0.546	Destabilizing
4	G22E	–2.031	Destabilizing	0.22	Stabilizing	–1.592	Destabilizing	–0.677	Destabilizing	–0.7547255	Stabilizing	–1.02	Destabilizing	0.83	Destabilizing	0.673	Stabilizing	0.213	Stabilizing
5	W25R	–0.932	Destabilizing	0.32	Stabilizing	–0.677	Destabilizing	–0.784	Destabilizing	–1.2714685	Destabilizing	–1.49	Destabilizing	0.87	Destabilizing	–0.286	Destabilizing	–0.759	Destabilizing
6	R33Q	–1.015	Destabilizing	–1.97	Destabilizing	–1.288	Destabilizing	–0.799	Destabilizing	–1.47021	Destabilizing	–0.51	Destabilizing	0.741	Destabilizing	–0.274	Destabilizing	–1.494	Destabilizing
7	L37P	–0.992	Destabilizing	–4.41	Destabilizing	–1.669	Destabilizing	–2.02	Destabilizing	–3.246535	Destabilizing	–1.78	Destabilizing	0.798	Destabilizing	–0.519	Destabilizing	–0.962	Destabilizing
8	P71T	–1.078	Destabilizing	0.16	Stabilizing	–0.704	Destabilizing	–1.311	Destabilizing	–1.0029515	Destabilizing	–1.56	Destabilizing	0.799	Destabilizing	0.284	Stabilizing	0.743	Stabilizing
9	G60C	–0.682	Destabilizing	–2.37	Destabilizing	–1.054	Destabilizing	0.064	Stabilizing	–1.683805	Destabilizing	–1.15	Destabilizing	0.53	Destabilizing	–0.085	Destabilizing	–1.339	Destabilizing
10	R76C	–0.995	Destabilizing	–0.12	Destabilizing	–0.827	Destabilizing	–0.963	Destabilizing	–0.2137374	Stabilizing	–1.12	Destabilizing	0.833	Destabilizing	–0.568	Destabilizing	–0.329	Destabilizing
11	R76H	–1.506	Destabilizing	–0.56	Destabilizing	–1.559	Destabilizing	–1.285	Destabilizing	–0.795757	Destabilizing	–1.74	Destabilizing	0.845	Destabilizing	–0.002	Destabilizing	–0.222	Destabilizing
12	V79F	–0.62	Destabilizing	–0.41	Destabilizing	–0.611	Destabilizing	–1.201	Destabilizing	–1.22595	Destabilizing	–1.39	Destabilizing	0.82	Destabilizing	–0.001	Destabilizing	–0.261	Destabilizing
13	F84C	–1.444	Destabilizing	–0.2	Destabilizing	–1.285	Destabilizing	–1.399	Destabilizing	–2.00914	Destabilizing	–1.97	Destabilizing	0.66	Destabilizing	–2.445	Destabilizing	–1.137	Destabilizing
14	V85M	–0.654	Destabilizing	–0.63	Destabilizing	–0.59	Destabilizing	–0.898	Destabilizing	–1.330545	Destabilizing	–1.67	Destabilizing	0.73	Destabilizing	0.206	Stabilizing	–0.056	Destabilizing
15	V85G	–2.703	Destabilizing	–2.37	Destabilizing	–3.095	Destabilizing	–2.23	Destabilizing	–3.778675	Destabilizing	–3.04	Destabilizing	0.83	Destabilizing	–0.887	Destabilizing	–2.258	Destabilizing
16	P88L	0.012	Stabilizing	2.12	Stabilizing	0.838	Stabilizing	0.36	Stabilizing	–0.1627065	Stabilizing	–0.86	Destabilizing	0.56	Destabilizing	0.321	Stabilizing	1.439	Stabilizing
17	Y98S	–1.096	Destabilizing	–0.54	Destabilizing	–0.902	Destabilizing	–1.147	Destabilizing	–1.290914	Destabilizing	–1.51	Destabilizing	0.74	Destabilizing	0.042	Stabilizing	0.412	Stabilizing
18	T154N	–1.13	Destabilizing	–0.15	Destabilizing	–0.847	Destabilizing	–1.141	Destabilizing	–1.261667	Destabilizing	–1.33	Destabilizing	0.76	Destabilizing	–0.25	Destabilizing	0.572	Stabilizing
19	Y177C	–1.423	Destabilizing	–0.87	Destabilizing	–1.336	Destabilizing	–1.443	Destabilizing	–1.922375	Destabilizing	–1.26	Destabilizing	0.81	Destabilizing	–1.43	Destabilizing	–0.921	Destabilizing
20	G178E	0.098	Stabilizing	–3.86	Destabilizing	–0.407	Destabilizing	–0.524	Destabilizing	–1.0358325	Destabilizing	–0.71	Destabilizing	0.76	Destabilizing	–0.061	Destabilizing	–1.608	Destabilizing
21	H194L	1.002	Stabilizing	0.21	Stabilizing	0.958	Stabilizing	0.235	Stabilizing	–0.206911	Stabilizing	0.6	Stabilizing	0.7	Stabilizing	–0.084	Destabilizing	0.359	Stabilizing
22	R202H	–1.236	Destabilizing	0.5	Stabilizing	–1.092	Destabilizing	–1.162	Destabilizing	–0.8287845	Destabilizing	–1.48	Destabilizing	0.73	Destabilizing	–0.114	Destabilizing	–0.362	Destabilizing
23	T204K	–0.521	Destabilizing	0.9	Stabilizing	0.107	Stabilizing	–0.918	Destabilizing	–0.5708325	Stabilizing	–1.12	Destabilizing	0.64	Stabilizing	–0.081	Destabilizing	0.041	Stabilizing
24	T204M	0.116	Stabilizing	0.62	Stabilizing	0.453	Stabilizing	–0.193	Destabilizing	–0.4244125	Stabilizing	0.02	Destabilizing	0.67	Stabilizing	–0.073	Destabilizing	0.3	Stabilizing
25	L214P	–1.075	Destabilizing	–3.14	Destabilizing	–1.508	Destabilizing	–2.303	Destabilizing	–3.150245	Destabilizing	–1.47	Destabilizing	0.83	Destabilizing	–0.66	Destabilizing	–1.133	Destabilizing
26	E205K	–0.457	Destabilizing	–1	Destabilizing	–0.282	Destabilizing	–1.634	Destabilizing	–0.4850575	Stabilizing	–0.68	Destabilizing	0.81	Destabilizing	0.026	Stabilizing	–0.302	Destabilizing
27	E227D	–1.289	Destabilizing	–1.99	Destabilizing	–1.477	Destabilizing	–0.869	Destabilizing	–1.095409	Destabilizing	–0.19	Destabilizing	0.605	Stabilizing	–0.477	Destabilizing	–1.291	Destabilizing
28	S18P	–0.521	Destabilizing	–0.45	Destabilizing	–0.415	Destabilizing	–1.236	Destabilizing	–0.9889335	Stabilizing	–0.01	Stabilizing	0.81	Stabilizing	0.013	Stabilizing	0.349	Stabilizing
29	G60S	–0.656	Destabilizing	–3.84	Destabilizing	–1.102	Destabilizing	–0.242	Destabilizing	–1.0329635	Destabilizing	–1.42	Destabilizing	0.53	Destabilizing	–0.082	Destabilizing	–1.37	Destabilizing
30	L11P	–0.728	Destabilizing	–2.63	Destabilizing	–1.015	Destabilizing	–1.95	Destabilizing	–2.72524	Destabilizing	–1.54	Destabilizing	0.909	Destabilizing	–0.427	Destabilizing	–0.531	Destabilizing
31	G22R	–0.935	Destabilizing	–0.46	Destabilizing	–0.673	Destabilizing	–0.704	Destabilizing	–0.365542	Stabilizing	–1.04	Destabilizing	0.79	Destabilizing	1.346	Stabilizing	1.374	Stabilizing

(*Contd.*)

Table 3 — Stability analysis of Cx43 selected missenses (*Contd.*)

				DUET				Mupro		INPS–MD		I MUTANT 2.0 SEQ				Dyna Mut			
		mCSM		SDM		DUET		Mupro		INPS–MD		I MUTANT 2.0 SEQ		I Stabilizing		ENCOM		Dyna Mut	
Sl. no	Mutations	ΔΔG	Stability	ΔΔG	Stability	ΔΔG	Stability	ΔΔG	Stability	ΔΔG	Stability	ΔΔG	Stability	ΔΔG	Stability	ΔΔG	Stability	ΔΔG	Stability
32	W25C	–1.106	Destabilizing	0.38	Stabilizing	–0.692	Destabilizing	–0.55	Destabilizing	–1.441255	Destabilizing	–2.1	Destabilizing	0.86	Destabilizing	–0.32	Destabilizing	–0.616	Destabilizing
33	D47H	–0.78	Destabilizing	0.53	Stabilizing	–0.6	Destabilizing	–0.671	Destabilizing	–0.596999	Stabilizing	–0.11	Stabilizing	0.51	Stabilizing	0.192	Stabilizing	0.567	Stabilizing
34	Q49P	–0.274	Destabilizing	–1.12	Destabilizing	–0.28	Destabilizing	–0.896	Destabilizing	–0.770067	Stabilizing	–0.41	Destabilizing	0.71	Destabilizing	–0.162	Destabilizing	–0.186	Destabilizing
35	Q58H	–0.495	Destabilizing	0.79	Stabilizing	–0.294	Destabilizing	–0.677	Destabilizing	–0.5396875	Stabilizing	–0.73	Destabilizing	0.808	Destabilizing	–0.018	Destabilizing	0.03	Stabilizing
36	P59H	–0.001	Destabilizing	0.16	Stabilizing	0.041	Stabilizing	–0.738	Destabilizing	–0.5029356	Stabilizing	–1.64	Destabilizing	0.86	Destabilizing	0.119	Stabilizing	0.276	Stabilizing
37	P59A	–0.239	Destabilizing	–0.2	Destabilizing	–0.06	Destabilizing	–0.852	Destabilizing	–0.7431565	Stabilizing	–1.86	Destabilizing	0.84	Destabilizing	0.117	Stabilizing	0.073	Stabilizing
38	S86Y	–0.358	Destabilizing	0.33	Stabilizing	–0.16	Destabilizing	–0.592	Destabilizing	–0.2277534	Stabilizing	0.08	Destabilizing	0.84	Destabilizing	0.744	Stabilizing	1.867	Stabilizing
39	V96E	–2.489	Destabilizing	–1.51	Destabilizing	–2.518	Destabilizing	–1.454	Destabilizing	–2.522035	Destabilizing	–1.86	Destabilizing	0.5	Stabilizing	–0.153	Destabilizing	–0.499	Destabilizing
40	Y98C	–0.161	Destabilizing	0.12	Stabilizing	0.121	Stabilizing	–0.885	Destabilizing	–1.150707	Destabilizing	–1.03	Destabilizing	0.7	Destabilizing	0.109	Stabilizing	0.393	Stabilizing
41	P193L	–0.872	Destabilizing	–0.07	Destabilizing	–0.599	Destabilizing	–0.126	Destabilizing	–0.9917315	Stabilizing	–0.73	Destabilizing	0.83	Destabilizing	0.088	Stabilizing	0.225	Stabilizing
42	H194P	0.854	Stabilizing	–0.32	Destabilizing	0.749	Stabilizing	–0.614	Destabilizing	–0.5732749	Stabilizing	0.42	Stabilizing	0.61	Destabilizing	0.461	Stabilizing	1.045	Stabilizing
43	S201Y	–0.41	Destabilizing	–0.28	Destabilizing	–0.418	Destabilizing	–0.803	Destabilizing	–0.5128665	Stabilizing	0.12	Stabilizing	0.58	Destabilizing	1.093	Stabilizing	1.378	Stabilizing
44	S201F	–0.636	Destabilizing	0.6	Stabilizing	–0.37	Destabilizing	–0.558	Destabilizing	–0.829358	Stabilizing	0.43	Stabilizing	0.61	Destabilizing	0.584	Stabilizing	0.83	Stabilizing
45	K206R	–0.934	Destabilizing	–0.78	Destabilizing	–0.769	Destabilizing	–0.384	Destabilizing	–0.9249555	Stabilizing	–0.14	Destabilizing	0.74	Destabilizing	0.129	Stabilizing	0.458	Stabilizing
46	S220Y	–0.656	Destabilizing	0.3	Stabilizing	–0.444	Destabilizing	0.123	Stabilizing	–0.378394	Stabilizing	–0.37	Destabilizing	0.7	Stabilizing	1.269	Stabilizing	1.86	Stabilizing

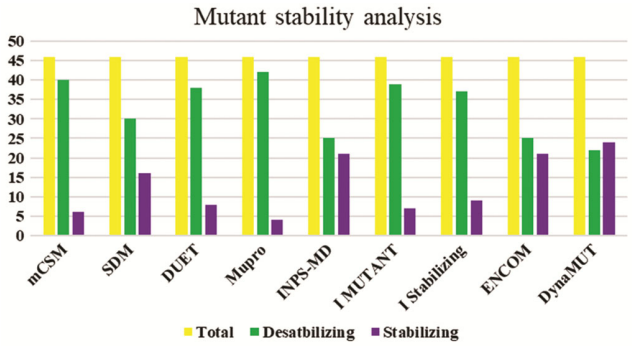


Fig. 3 — Server-based mutation stability analysis in Cx43 protein

Conservation analysis of Cx43

ConSurf server was used to analyze the conservative sites in the Cx43 protein. A mutation L214P was screened for conservative analysis, and the conSurf result revealed that the L214 position is subject to have more conserved at a scale of 7 (Fig. 4). From this analysis, L214P mutation was testimony for modelling and was a potent inhibitor analysis.

Structure modelling and validation

A crystal structure of Cx43 was not completely available in RCSB PDB; thus, a sequence of the Cx43 gene was retrieved from the Uniport database (Uniport ID: P17302) and submitted to I-TASSER web-based server. It generally retrieves template structure from the RCSB PDB library based on similar folds via a threading approach, and I-TASSER then utilizes the SPICKER program to cluster the confirmations through pairwise sequence alignment (PSA). As a result, five models generated with a confidence score from that model 1 with the best score were selected for further analysis. Cx43 was mutated by replacing the LEU at the 214th position

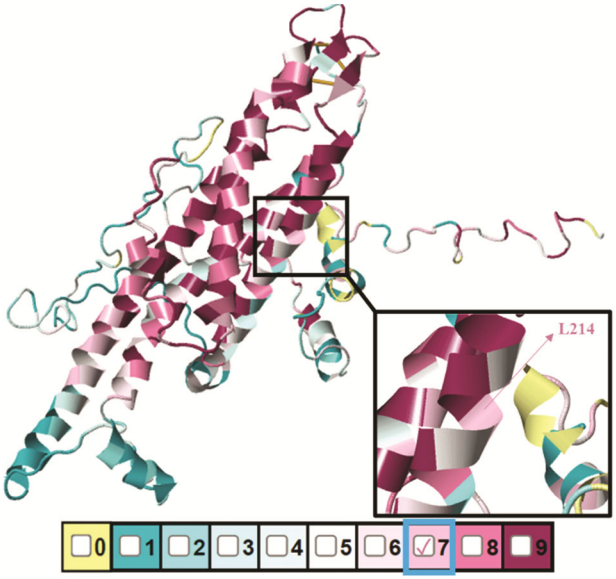


Fig. 4 — Conservative analysis of LEU at 214th position in native Cx43 protein

with PRO and submitted in I-TASSER (Fig. 5A & B). The native and mutant Cx43 structure was validated by the Ramachandran plot server, which obtained 94.97% and 94.362% of residues distributed in the highly preferred region of favored regions (Fig. 6).

Molecular docking and inhibitor analysis

In the present study, 36 compounds of interest docked with L214P mutated Cx43 protein (Table 4), which revealed the compounds Kanamycin, Ginsenoside, and Astragaloside IV shown to interact with mutated Cx43 with a maximum of 5 hydrogen bonds (Fig. 7A-C). The residues involved in the interaction are TYR155, GLY22, SER27, ASN302, ASN300, ARG293, ASN309, ARG148, LYS13 and

TYR286. Glycyrrhetinic acid, Halothane, Heptanol, Ketamine, Propofol, Quinine, pentachlorophenol, Rutaecarpine, Ascorbic acid 6-palmitate, Boldine, and Terbinafine doesn't have any hydrogen bond interactions. Other compounds showed an interaction between 1 to 4 hydrogen bonds.

AMDE result analysis revealed that high-affinity compounds Kanamycin and Astragaloside IV violated from Lipinski rule by three violations (Table 4). Ginsenoside has no violations and has high-affinity interactions of 5 hydrogen bonds with mutated Cx43 (Fig. 7B). Thus, Ginsenoside would be

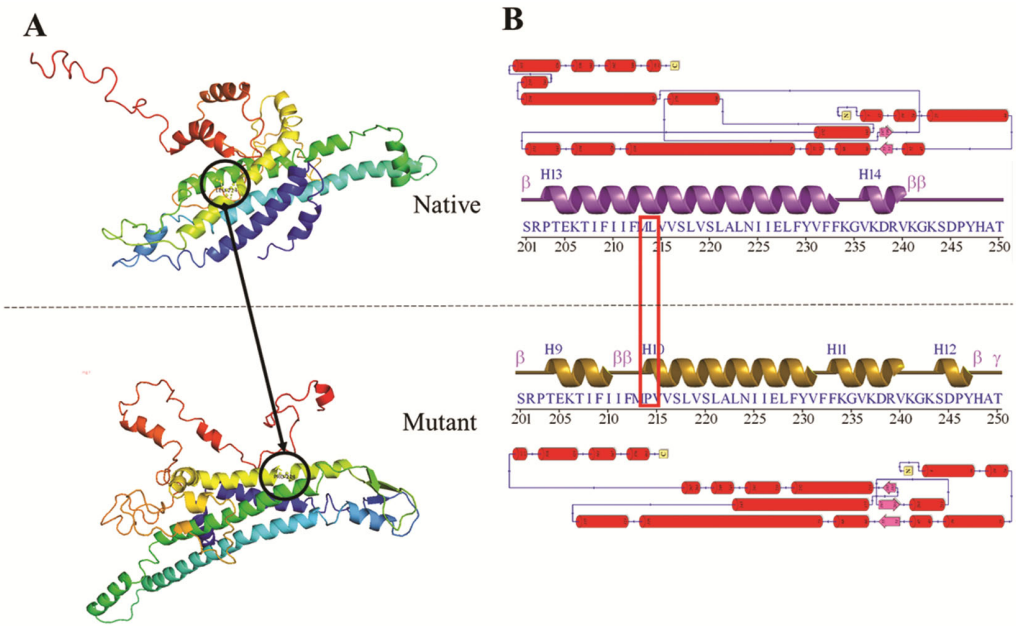


Fig. 5 — (A) 3D structure of native and mutant Cx43 protein; and (B) Secondary structural confirmation of LEU replaced with PRO at 214th position

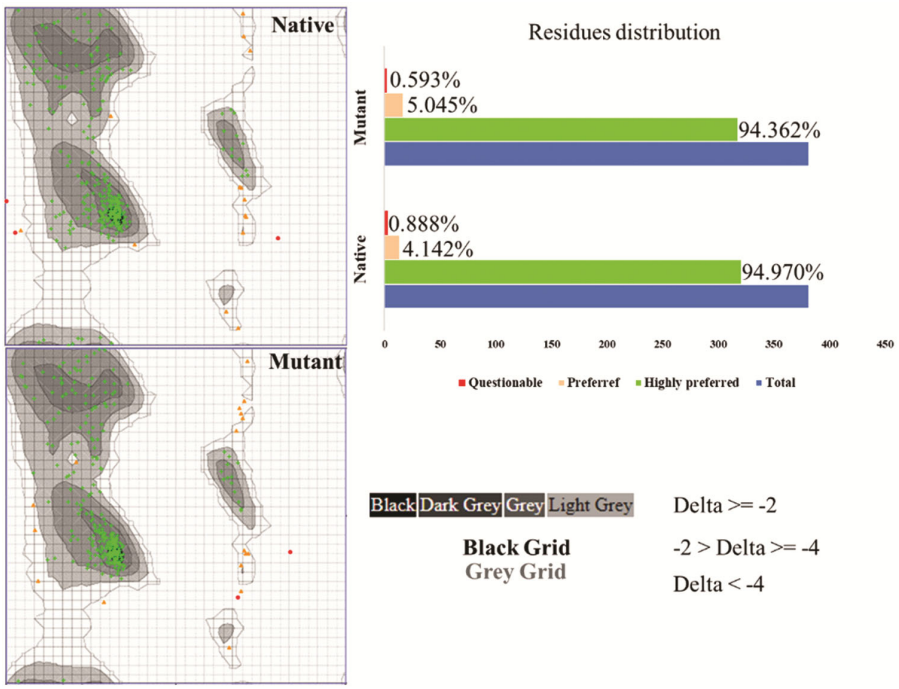
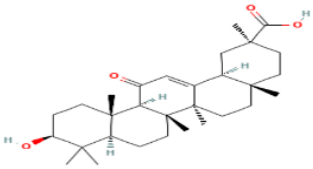
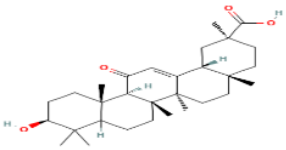
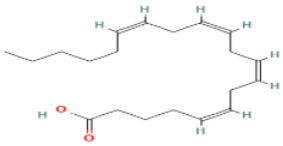
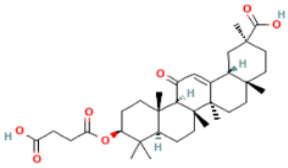
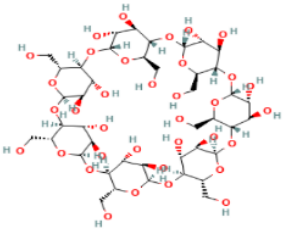
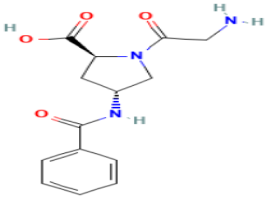
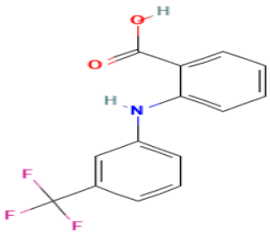
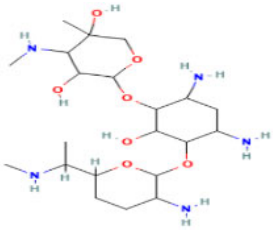


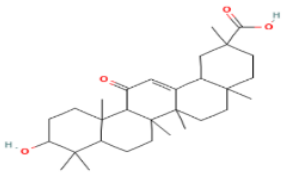
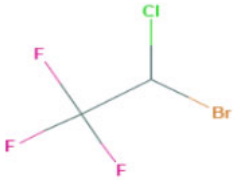
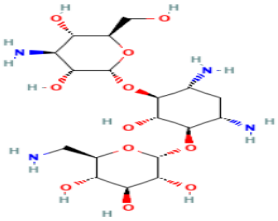
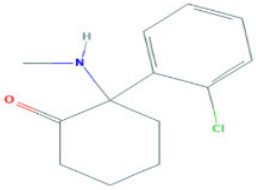
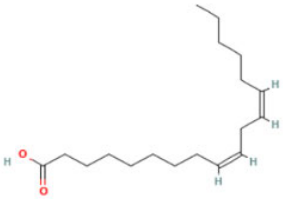
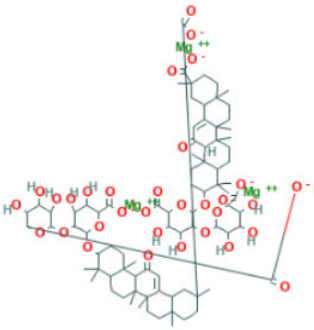
Fig. 6 — Ramachandran plot analysis of native and mutant Cx43 structure with residues distribution

Table 4 — Docking analysis and Lipinski score

Sl. no	Inhibitors	Class	Structure	PCID	MW (g/mol)	Binding energy (kCal/mole)	H Bonds	Lipinski violation
1	18α-glycyrrhetinic acid	Bioactive compound derivative		73398	470.7	–9.5	2	1 (MLOGP>4.15)
2	18β-glycyrrhetinic acid	Bioactive compound derivative		44435791	470.7	–9.8	1	1 (MLOGP>4.15)
4	Arachidonic acid	Fatty acid		444899	304.5	–5.5	1	1 (MLOGP>4.15)
5	Carbenoxolone	Chemical compound		636403	570.8	–8.9	4	2 (MW>500, MLOGP>4.15)
6	Cyclodextrins	Polysaccharide		444041	1135.0	–6.9	4	3 (MW>500, N or O>10, NH or OH>5)
7	Danegaptide	Peptide		16656685	291.30	–6.6	4	0
8	Flufenamic acid	NSAID		3371	281.23	–7.9	2	0
9	Gentamicin	Antibiotic		3467	477.6	–7.1	4	2 (N or O>10, NH or OH>5)

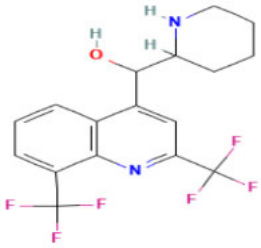
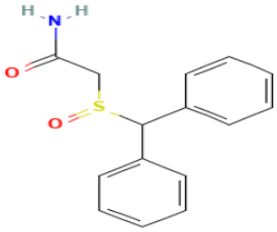
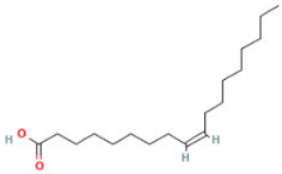
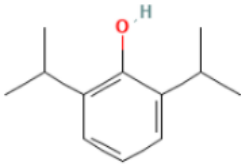
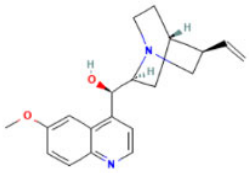
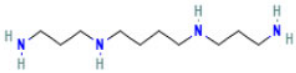
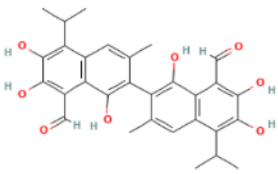
(*Contd.*)

Table 4 — Docking analysis and Lipinski score (*Contd.*)

Sl. no	Inhibitors	Class	Structure	PCID	MW (g/mol)	Binding energy (kCal/mole)	H Bonds	Lipinski violation
10	Glycyrrhetinic acid	Bioactive compound		3230	470.7	–9.8	0	1 (MLOGP>4.15)
11	Halothane	Anesthetic		3562	197.38	–4.4	0	0
12	Heptanol	Chemical compound		8129	116.20	–4.1	0	0
13	Kanamycin	Antibiotic		6032	484.5	–6.0	5	2 (N or O>10, NH or OH>5)
14	Ketamine	Anesthetic		3821	237.72	–6.9	0	0
15	Linoleic acid	Fatty acid		5280450	280.4	–5.5	1	1 (MLOGP>4.15)
16	Magnesium isoglycyrrhizinate	Synthesized bioactive compound		139032961	1712.7	–9.5	3	NA

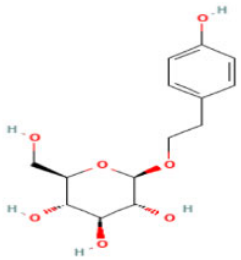
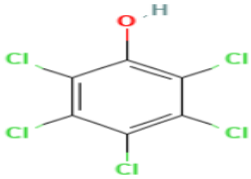
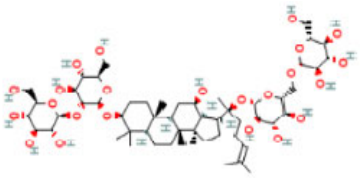
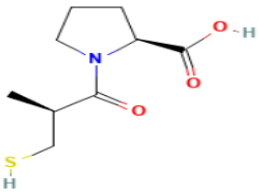
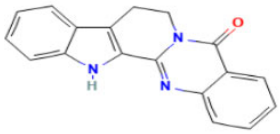
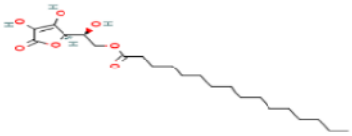
(*Contd.*)

Table 4 — Docking analysis and Lipinski score (*Contd.*)

Sl. no	Inhibitors	Class	Structure	PCID	MW (g/mol)	Binding energy (kCal/mole)	H Bonds	Lipinski violation
17	Mefloquine	Quinolines		4046	378.31	–8.7	2	0
18	Modafinil	Chemical compound		4236	273.4	–5.1	2	0
19	Octanol	Chemical compound		957	130.23	–4.5	1	0
20	Oleic acid	Fatty acid		445639	282.5	–4.6	1	1 (MLOGP>4.15)
21	Propofol	Chemical compound		4943	178.27	–6.2	0	0
22	Quinine	Quinolines		3034034	324.4	–8.1	0	0
23	Spermine	Polyamines		1103	202.34	–4.0	1	0
24	Gossypol	Bioactive compound		3503	518.6	–8.5	3	2 (MW>500, NH or OH>5)

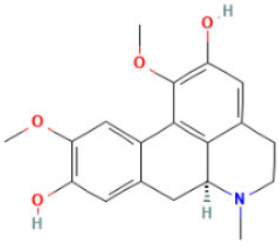
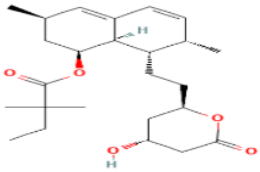
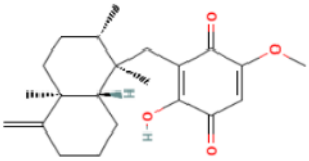
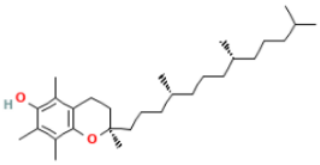
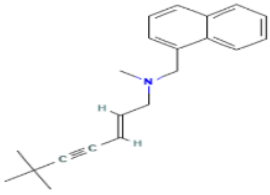
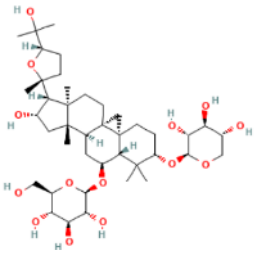
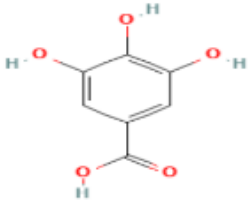
(*Contd.*)

Table 4 — Docking analysis and Lipinski score (*Contd.*)

Sl. no	Inhibitors	Class	Structure	PCID	MW (g/mol)	Binding energy (kCal/mole)	H Bonds	Lipinski violation
25	Salidroside	Glycosides Compound		159278	300.30	–6.7	3	0
26	Pentachlorophenol	Chemical compound		992	266.3	–5.5	0	1 (MLOGP>4.15
27	Ginsenoside	Steroids		3086007	444.7	–8.5	5	1 (MLOGP>4.15)
28	Captopril	Chemical compound		44093	217.29	–8.0	1	0
29	Rutaecarpine	Chemical compound		65752	287.3	–8.7	0	0
30	Ascorbic acid 6-palmitate	Fatty acid		54680660	414.5	–6.2	0	0

(*Contd.*)

Table 4 — Docking analysis and Lipinski score (*Contd.*)

Sl. no	Inhibitors	Class	Structure	PCID	MW (g/mol)	Binding energy (kCal/mole)	H Bonds	Lipinski violation
31	Boldine	Alkaloid		10154	327.4	–6.5	0	0
32	Simvastatin	Synthesized bioactive compound		54454	418.6	–8.8	1	0
33	Ilimaquinone	Quinolines		72291	358.5	–8.4	1	0
34	α-tocopherol	Vitamins		14985	430.7	–6.0	1	1 (MLOGP>4.15)
35	Terbinafine	Allylamine derivative		1549008	291.4	–8.0	0	1 (MLOGP>4.15)
36	Astragaloside IV	Triterpenoid		13943297	785.0	–8.9	5	3 (MW>500, N or O>10, NH or OH>5)
37	Gallic acid	Bioactive compound		370	170.12	–5.3	4	0

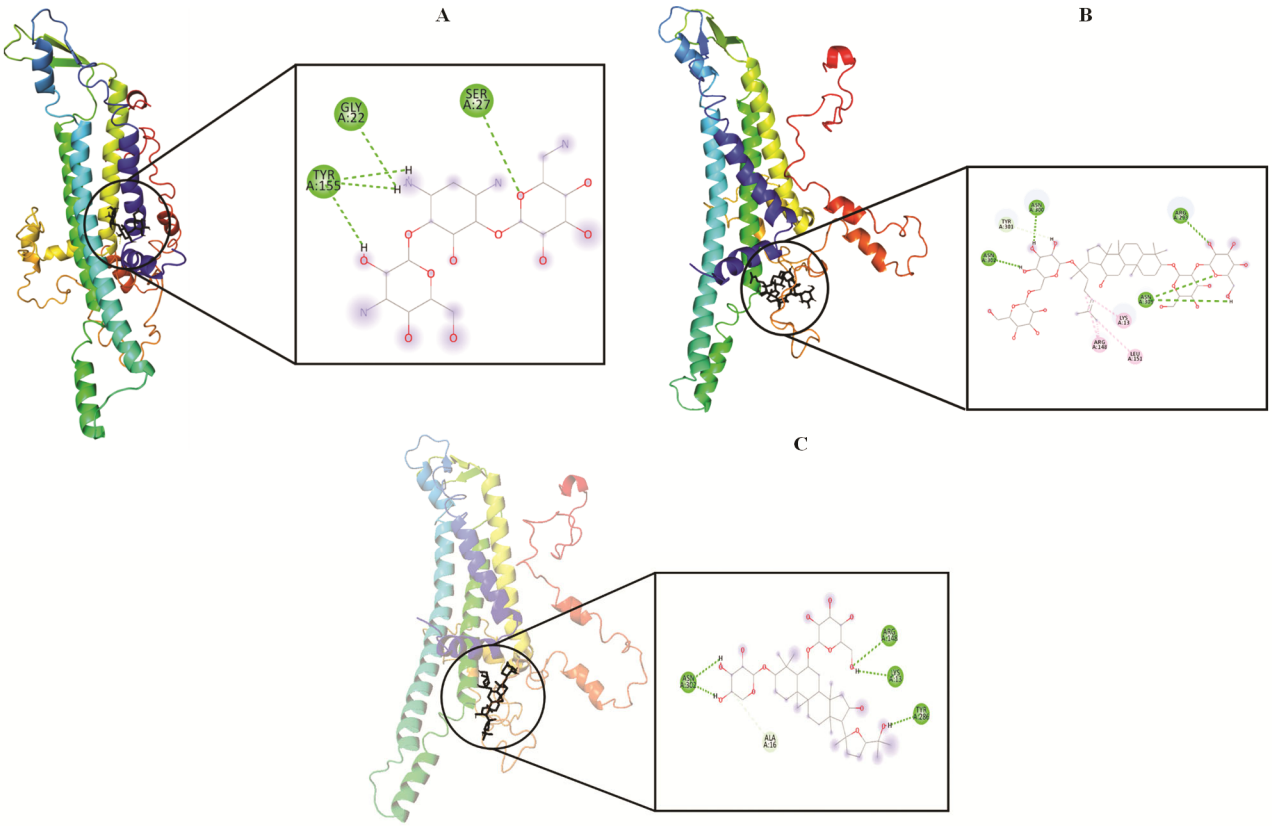


Fig. 7 — Docking analysis of (A) Kanamycin; (B) Ginsenoside; and (C) Astragaloside IV with mutated Cx43

the better compound to inhibit the L214P mutated Cx43 protein.

Discussion

In the present study, we collected missense variants of Cx43 from different databases and literature and then identified the pathogenic mutations using five different algorithms from 219 variants. The prediction from the pathogenicity determinations server has 52 as deleterious (Table 1). Pathogenic mutations are often proven to affect protein function[22]. The deleterious mutations are analyzed for functional impact using a mutation assessor prediction server from those 46 mutations found to impact protein functions (Table 2). The protein stability of mutations was further confirmed by nine different servers, which obtained eight destabilizing mutations (Table 3). Impacts in the protein functions are primarily due to the destabilization of the protein structure[23-25].

Further, these mutations were compared with the binding pocket of Cx43, which shows L214 has been observed in the binding pocket. Which was then analyzed for a conserved position that revealed L214 kepta significant position in structural changes (Figs 4 & 6). The studies reported that an amino acid substitution at the ligand-binding site significantly alters the ligand specificity and binding affinity. Thus, inhibitors with the best binding affinity, even at mutated conditions, are important[26]. With the evidence of pathogenicity, functional impact, and structural changes, a Cx43 protein was mutated with L214P and analyzed for a potent inhibitor.

Several studies have proven that mutation at the atomic level has a severe impact on structural changes, stability, and functions of the protein[14,27].

A comparative computational approach anticipated the effect of disease-mediating missense variants in the protein structural and functional impacts[28]. In our study, the mutations such as R76H; V79F; F84C; V85G; Y177C; L214P; G60S; L11P are obtained as disease-causing and structural changes mutations in Cx43. Thus, these insights let us understand the genotype-phenotype correlation of genetic diseases related to Cx43 and assisted in scrutinizing the prioritized pathogenic mutations[29]. Based on the structural stability and binding pocket analysis, it was found that mutations at the binding pocket result to be a significant structural change. As

evidence, in (Fig. 6B), a secondary Cx43 structure (red pipeline diagram) shows two beta-strands, 18 helices, 26 helix-helix interactions, 42 beta turns, and 14 gamma turns. Three disulfides from native Cx43 were changed to 3 beta-strands, 17 helices, 19 helix-helix interactions, 50 beta turns, 23 gamma turns, and three disulfides. Research also reported that single point mutation leads to protein-misfolding or structural changes and aggregation, which is the primary cause of various diseases[30-32]. Because of mutations in Cx43, which cause various diseases, the need for drugs targeted for mutated Cx43 is recommended. A computational approach is one of the cost-efficient, time-saving and scrutinizing platforms in the field of drug discovery due to hungry for unravelling the drugs for targeted mutations in various diseases, predominantly genetic disorders.

Several studies have aimed to discover a potent drug to inhibit Cx43 protein, which Refs reports[4,5] is also depicted in (Table 4). However, no studies have yet to corroborate the inhibitors for mutated Cx43, and we aimed to analyze the interactions of Cx43 inhibitors with the native Cx43 (data not shown) and L214P mutant Cx43 protein. This analysis was obtained using the list of Cx43 inhibitors retrieved from the literature survey (Table 4). After that, Autodock vina software was used to do a virtual screening analysis on 36 Cx43 inhibitors. Among them, 30 followed the Lipinski rule of 5, from which the compound ginsenoside showed the strongest affinity (hydrogen bond: 5; binding energy: -8.5 kcal/mol) with L214P mutant Cx43 protein compared to other inhibitors (Table 4 and Fig. 7b).

Ginsenoside is a primary active compound of *Panax ginseng*, a Korean traditional medicine for longevity. There are numerous clinical studies have been conducted on various chronic diseases[33]. Also, it has been reported that ginsenosides can bind with targeted proteins in the cells, leading to beneficial effects[34]. A study reported that ginsenosides downregulated the expression of Cx43 in Bisphenol A-induced testicular toxicity[35]. In our research, ginsenosides interacted efficiently with the L214P mutant Cx43 protein. Thus, it might be a better inhibitor of native Cx43 and mutated Cx43 with a potential drug as personalized medicine.

Conclusion

This is the first study reporting that substituting leucine at the 214th position with proline could be the most pathogenic mutation in disease-causing role in Cx43 protein based on the computational method. Pathogenicity of the variant was confirmed by deleterious, functional, and structural assessment of mutations. A COACH-D and CornSurf server results revealed that a residue LEU 214 significantly participated in ligand binding sites and was the most conserved residue. Further, a structure-modelled mutant with the desired variation was observed as the entire protein structure changed (Fig. 6b), which was then performed molecular docking analysis to screen potent inhibitors. The compound Kanamycin, Ginsenoside, and Astragaloside IV are better interactions with Cx43 mutants with a maximum of 5 hydrogen bonds. Ginsenoside is the only compound that follows a Lipinski rule of five. Thus, the result obtained from this study suggests that Ginsenoside would be a better potent inhibitor for native and mutant Cx43 in most genetic diseases and could therefore be a candidate for personalized medicine.

Acknowledgement

The authors are grateful to the Vellore Institute of Technology, Vellore, for providing the facility and opportunity to carry out this work.

Conflict of interest

All authors declare no conflict of interest.

References

1 Sinyuk M, Mulkearns-Hubert EE, Reizes O & Lathia J, Cancer Connectors: Connexins, Gap Junctions, and Communication. *Front Oncol*, 8 (2018) 646.
2 Burendei B, Shinozaki R, Watanabe M, Terada T, Tani K & Fujiyoshi Y, Cryo-EM structures of undocked innexin-6 hemichannels in phospholipids. *Sci Adv*, 6 (2020) 3157.
3 Srinivas M, Verselis VK & White TW, Human diseases associated with connexin mutations. *Biochim Biophys Acta*, 1860 (2018) 192.
4 Natha CM, Vemulapalli V, Fiori MC, Chang CWT & Altenberg GA, Connexin hemichannel inhibitors with a focus on aminoglycosides. *Biochim Biophys Acta Mol Basis Dis*, 1867 (2021) 166115.
5 Katturajan R & Price SE, A role of connexin 43 on the drug-induced liver, kidney, and gastrointestinal tract toxicity with associated signaling pathways. *Life Sci*, 280 (2021) 119629.
6 Ishida-Yamamoto A, Erythrokeratodermia variabilis et progressiva. *J Dermatol*, 43 (2016) 280.
7 Umegaki-Arao N, Sasaki T, Fujita H, Aoki S, Kameyama K & Amagai M, Inflammatory Linear Verrucous Epidermal Nevus with a Postzygotic GJA1 Mutation Is a Mosaic Erythrokeratodermia Variabilis et Progressiva. *J Invest Dermatol*, 137 (2017) 967.
8 Wang H, Cao X, Lin Z, Lee M, Jia X & Ren Y, Exome sequencing reveals mutation in GJA1 as a cause of keratoderma-hypotrichosis-leukonychia totalis syndrome. *Hum Mol Genet*, 24 (2015) 243.

9 García IE, Prado P, Pupo A, Jara O, Rojas-Gómez D & Mujica P,Connexinopathies: a structural and functional glimpse. *BMC Cell Biol*, 17 (2016) 17.

10 Gago-Fuentes R, Fernández-Puente P, Megias D, Carpintero-Fernández P, Mateos J & Acea B, Proteomic Analysis of Connexin 43 Reveals Novel Interactors Related to Osteoarthritis. *Mol Cell Proteomics MCP*, 14 (2015) 1831.

11 Panchal NK, Bhale A, Verma VK & Beevi SS, Computational and molecular dynamics simulation approach to analyze the impactof XPD gene mutation on protein stability and function. *Mol Simul*, 46 (2020) 1200.

12 Rangasamy N, Kumar NS & Santhy KS, Computational analysis of missense variants in MMP2 gene linked with Winchester syndrome and Nodulosis-Arthropathy-Osteolysis reveals structural shift in protein-protein and protein-ligand complexes. *Meta Gene*, 29 (2021) 100931.

13 Mosaeilhy A, Mohamed MM, C GPD, El Abd HSA, Gamal R & Zaki OK, Genotype-phenotype correlation in 18 Egyptian patients with glutaric acidemia type I. *Metab Brain Dis*, 32 (2017) 417.

14 Thirumal Kumar D, Jain N, Evangeline J, Kamaraj B, Siva R & Zayed H, A computational approach for investigating the mutational landscape of RAC-alpha serine/threonine-protein kinase (AKT1) and screening inhibitors against the oncogenic E17K mutation causing breast cancer. *Comput Biol Med*, 115 (2019) 103513.

15 Katsonis P, Koire A, Wilson SJ, Hsu TK, Lua RC & Wilkins AD, Single nucleotide variations: Biological impact and theoretical interpretation. *Protein Sci Publ Protein Soc*, 23 (2014) 1650.

16 Deller MC, Kong L & Rupp B, Protein stability: a crystallographer's perspective. *Acta Crystallogr Sect F Struct Biol Commun*, 72 (2016) 72.

17 Rodrigues CH, Pires DE & Ascher DB, DynaMut: predicting the impact of mutations on protein conformation, flexibility and stability. *Nucleic Acids Res*, 2018.

18 Wu Q, Peng Z, Zhang Y & Yang J, COACH-D: improved protein–ligand binding sites prediction with refined ligand-binding poses through molecular docking. *Nucleic Acids Res*, 46 (2018) 438.

19 Honorato RV, Koukos PI, Jiménez-García B, Tsaregorodtsev A, Verlato M & Giachetti A, Structural Biology in the Clouds: The WeNMR-EOSC Ecosystem. *Front Mol Biosci*, 8 (2021) 708.

20 Ben Chorin A, Masrati G, Kessel A, Narunsky A, Sprinzak J & Lahav S, ConSurf-DB: An accessible repository for the evolutionary conservation patterns of the majority of PDB proteins. *Protein Sci Publ Protein Soc*, 29 (2020) 258.

21 Ha EJ, Lwin CT & Durrant JD, LigGrep: a tool for filtering docked poses to improve virtual-screening hit rates. *J Cheminformatics*, 12 (2020) 69.

22 Sun H & Yu G, New insights into the pathogenicity of nonsynonymous variants through multi-level analysis. *Sci Rep*, 9 (2019) 1667.

23 Shi Z & Moult J, Structural and Functional Impact of Cancer Related Missense Somatic Mutations. *J Mol Biol*, 413 (2011) 495.

24 Saffari-Chaleshtori J, Mohammad Shafiee S & Heidarian E, The effect of bilirubin on Bad, Bak, and Bim pro-apoptotic factors: A molecular dynamic simulation study. *Indian J Biochem Biophys*, 58 (2021) 236.

25 Agrawal A, Awasthi R & Kulkarni GT, A bioinformatic approach to establish P38α MAPK inhibitory mechanism of selected natural products in psoriasis. *Indian J Biochem Biophys*, 59 (2022) 165.

26 Ricatti J, Acquasaliente L, Ribaudo G, De Filippis V, Bellini M, Llovera RE, Barollo S, Pezzani R, Zagotto G, Persaud KC & Mucignat-Caretta C, Effects of point mutations in the binding pocket of the mouse major urinary protein MUP20 on ligand affinity and specificity. *Sci Rep*, 9 (2019) 300.

27 Kumar A & Purohit R, Cancer Associated E17K Mutation Causes Rapid Conformational Drift in AKT1 Pleckstrin Homology (PH) Domain. *PLoS One*, 8 (2013) 64364.

28 Agrahari AK, Sneha P, George Priya Doss C, Siva R & Zayed H, A profound computational study to prioritize the disease-causing mutations in PRPS1 gene. *Metab Brain Dis*, 33 (2018) 589.

29 Tanwar H, Kumar DT, Doss CGP & Zayed H, Bioinformatics classification of mutations in patients with Mucopolysaccharidosis IIIA. *Metab Brain Dis*, 34 (2019) 1577.

30 Wang F, Orioli S, Ianeselli A, Spagnolli G, A Beccara S & Gershenson A, All-Atom Simulations Reveal How Single-Point Mutations Promote Serpin Misfolding. *Biophys J*, 114 (2018) 2083.

31 Singh P, Yadav M, Niveria K & Verma AK, Versatility of berberine as an effective immunomodulator and chemo sensitizer against p53 mutant cell. *Indian J Biochem Biophys*, 59 (2022) 509.

32 Janani DM & Usha B, *In silico* analysis of functional non-synonymous and intronic variants found in a polycystic ovarian syndrome (PCOS) candidate gene: DENND1A. *Indian J Biochem Biophys*, 57 (2020) 584.

33 Yu SE, Mwesige B, Yi YS & Yoo BC, Ginsenosides: the need to move forward from bench to clinical trials. *J Ginseng Res*, 43 (2019) 361.

34 Kim KH, Lee D, Lee HL, Kim CE, Jung K & Kang KS, Beneficial effects of Panax ginseng for the treatment and prevention of neurodegenerative diseases: past findings and future directions. *J Ginseng Res*, 42 (2018) 23.

35 Wang L, Hao J, Hu J, Pu J, Lü Z & Zhao L, Protective Effects of Ginsenosides against Bisphenol A-Induced Cytotoxicity in 15P-1 Sertoli Cells via Extracellular Signal-Regulated Kinase 1/2 Signalling and Antioxidant Mechanisms. *Basic Clin Pharmacol Toxicol*, 111 (2012) 42.