In-silico analysis of non-synonymous single nucleotide polymorphisms in human β-defensin type 1 gene reveals their impact on protein-ligand binding sites

Article *in* Computational Biology and Chemistry · March 2022

DOI: 10.1016/j.compbiolchem.2022.107669

CITATIONS
3

READS
84

5 authors, including:

Sana Ullah ..
Government College University Faisalabad
10 PUBLICATIONS **39** CITATIONS
SEE PROFILE

Todd Johnsen
ITMO University
1 PUBLICATION **3** CITATIONS
SEE PROFILE

Sergey Shityakov
University of Wuerzburg
165 PUBLICATIONS **1,841** CITATIONS
SEE PROFILE

Computational Biology and Chemistry 98 (2022) 107669

Contents lists available at ScienceDirect

Computational Biology and Chemistry

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

In-silico analysis of non-synonymous single nucleotide polymorphisms in human β-defensin type 1 gene reveals their impact on protein-ligand binding sites

Muhammad Mazhar Fareed [a,1], Sana Ullah [b,1], Shan Aziz [b], Todd Axel Johnsen [c], Sergey Shityakov [c,*]

[a] School of Science and Engineering, Department of Computer Science, Università degli studi di Verona, Verona, Italy
[b] Faculty of Life Sciences, Department of Bioinformatics and Biotechnology, Government College University, Faisalabad, Pakistan
[c] Laboratory of Chemoinformatics, Infochemistry Scientific Center, ITMO University, Saint-Petersburg, Russian Federation

ARTICLE INFO

Keywords:
In-silico
NsSNPs
Molecular
Human β-defensin
Single nucleotide polymorphism
Protein-ligand binding sites
Molecular docking
Molecular dynamics simulation

ABSTRACT

Single nucleotide polymorphism (SNPs) is an important genetic biomarker to assess protein function and its possible contribution to genetic diseases, such as the β- defensin 1 gene (DEFB1)-associated non-synonymous SNPs (nsSNPs). Defensins are antimicrobial and immunomodulatory peptides, acting as part of innate immunity, and killing bacteria by interacting phosphatidylinositol 4,5-bisphosphate (PIP2). Therefore, we apply cutting-edge computational algorithms to identify detrimental SNPs in the DEFB1 gene that potentially impact PIP2 binding sites. Furthermore, 4 most important nsSNPs in the DEFB1 gene were discovered (C67S, T58S, G62W, and Y35C) and only two of them were found to be linked to the PIP2 binding site-forming residues (Thr58 and Tyr35). Additional molecular docking and molecular dynamics simulations confirmed the decreased binding affinity of DEFB1 to bacterial PIP2 due to these mutations. Overall, this computational study analyzing nsSNPs in DEFB1 provides more understanding of how these missense mutations could impair or change protein functions by altering the PIP2 binding site.

1. Introduction

It is well known that single nucleotide polymorphisms (SNPs) might affect protein folding, stability, and function due to a change of the encoded amino acids in the protein structure (Shastry, 2009). It is roughly estimated that there are more than a million SNPs, either in coding and non-coding DNA regions (Herraez et al., 2009). In particular, the so-called non-synonymous SNPs (nsSNPs) are the point or missense mutations associated with changes in protein sequences and various pathologies. On the other hand, several *in silico* techniques have been developed to predict the nsSNP mutation probability due to its detrimental impact on protein function (Meade and O'Farrelly, 2019). These methods and tools were designed to reduce the experimental costs, speed, and SNP predictive accuracy for any protein of interest (Heimlich et al., 2014). In general, SNPs can be seen as surrogate pharmacogenomics biomarkers to test for association between a phenotype and a functional variant directly (Alwi, 2005). Indeed, using population-wide analysis of nsSNPs present in immunity-linked proteins can be useful to develop new medications and understand immune responses. In particular, SNP effects on innate immunity factors, such as defensins, are important to determine in conjunction with their antimicrobial activities (Machado and Ottolini, 2015). Human beta-defensins are small cysteine-rich cationic proteins exhibiting antimicrobial activity against a wide range of pathogens, such as viruses, bacteria, protozoa, and fungi (Lehrer et al., 2005). These molecules are essential for maintaining physiological homeostasis and serve as a primary barrier against oral, rhino, cutaneous, gastrointestinal, and genitourinary infections (Prado-Montes de Oca, 2010). Their antimicrobial action is associated with the cationic and amphipathic molecules, allowing them to bind to some components of bacterial membranes, such as phosphatidylinositol 4, 5-bisphosphate (PIP2). The distinct PIP2 binding sites were revealed in human defensin 2, which can be ablated by mutations. This could be

* Corresponding author.
E-mail addresses: mmazharfareed22@gcuf.edu.pk (M.M. Fareed), sanaullahlehre@gcuf.edu.pk (S. Ullah), shan.aziz@tuf.edu.pk (S. Aziz), todd.johnsen@infochemistry.ru (T.A. Johnsen), shityakoff@hotmail.com, shityakoff@itmo.ru (S. Shityakov).
[1] Joint first authors equally contributed.

https://doi.org/10.1016/j.compbiolchem.2022.107669
Received 7 February 2022; Received in revised form 18 March 2022; Accepted 19 March 2022
Available online 26 March 2022

Table 1
List of deleterious nsSNPs determined for human DEFB1.

Mutation	Poly-Phen score	PROVEAN score	SNP effect	
			SNP-PHD	SNP& GO
C67S	0.99	-9.69	Disease cause	Disease -effect
T58S	0.96	-3.87	Neutral	Neutral -effect
G62W	0.99	-6.12	Disease cause	Neutral -effect
Y35C	0.98	-8.63	Disease cause	Disease effect

Table 2
DEFB1 protein stability.

Mutation	ΔΔG, kcal/mol	ΔE^{ros}, REU	RMSD, Å
C67S	-1.18	-0.88	0.41
T58S	-0.99	-0.57	0.35
G62W	-0.87	4.14	0.8
Y35C	-1.15	-0.27	0.78

demonstrated by nsSNPs, neutralizing the defensin permeabilization effects on *Candida albicans* (Jarva et al., 2018). However, the structural impact of nsSNPs on defensin-ligand binding sites has yet to be determined. On the other hand, various computational and experimental data highlight the possibility of nsSNPs to alter human β-defensin 1 (DEFB1) function and stability. Nevertheless, they are lacking information about the PIP2 binding sites involved in this process (Subbiah et al., 2020; Jurevic et al., 2003). Therefore, in the current study, we wanted to apply various bioinformatics tools to analyze the human DEFB1 gene on possible nsSNPs. Additionally, there was also an interest in exploring how they might affect the DEFB1 activity to bind PIP2 in order to protect the human body against bacterial infection.

2. Computational methods

2.1. Data mining

The sequence of the human DEFB1 gene was obtained through the NCBI-SNP (only missense selected) database and the protein sequence was retrieved from the Uniprot database using the accession number: P60022 (Sherry et al., 2001; Consortium, 2015).

2.2. Identification of deleterious nsSNPs of DEFB1

For the prediction of missense mutational deleterious SNPs, several tools were implemented, such as the SIFT tool with the threshold 0.05 (Kumar et al., 2009), PolyPhen-v2 tool (Ghosh et al., 2021) for the identification of the key amino acids in target sequences with the scores ranging from 0.00 to 1.00. After the analysis by the SIFT and PloyPhenV2 algorithms to determine any deleterious nsSNPs in DEFB1 the PROVEAN program was used to predict whether an amino acid substitution has an impact on the biological function of the analyzed protein (Choi and Chan, 2015; Sim et al., 2012) using the scoring function of − 2.5 and below. The SNP variants were predicted with the implementation of the Predictor of Deleterious Single Nucleotide Polymorphisms tool (Adzhubei et al., 2013) and the Predicting Human Disease-related Mutations in Proteins with Functional Annotations tool integrated into the Snps&GO server (Abbas et al., 2019).

2.3. Protein stability prediction

For the identification of protein stability of DEFB1, we used the I-Mutant suite with a free energy change value of 3.0 kcal/mol. The wild-type DEFB1 protein structure was predicted with the AlphaFold algorithm (Jumper et al., 2021). The Rosetta refinement protocol, implementing the Monte-Carlo approach, was used to predict the protein stability via matching the root-mean-square-deviation (RMSD) values with the Rosetta energy parameters (Ramelot et al., 2009). The Rosetta energy differences for mutated forms (ΔE^{ros}_{mut}) were calculated using the following equation:

$$\Delta E^{ros}_{mut} = E^{ros}_{mut} - E^{ros}_{wt}$$

where E^{ros}_{wt} and E^{ros}_{mut} are the Rosetta energy values for the wild type and mutated forms, respectively.

2.4. Conservation analysis of nsSNPs

In the DEFB1 proteins affected by nsSNPs, the analysis of the evolutionary changes of amino acids was performed by the ConSurf web tool (Glaser et al., 2003; Shityakov et al., 2020). It identifies the rate of change in the amino acid in protein through the ConSurf calculation as a position of the amino acid; the evolutionary relationship along with homologous sequences and scores range between 1 and 9 (Glaser et al., 2003).

2.5. Post-translational modification prediction

The structural, functional, and mechanistic characterization of several types of post-translational modifications (PTMs), were predicted for the DEFB1 protein by using the ModPred tool (Pejaver et al., 2014). This web tool can predict up to 23 types of post-translational modifications in DEFB1.

2.6. Gene association network analysis

To find other genes that are related to a set of the DEFB1 gene, using a very large set of functional association data, we used the GeneMANIA tool (Mostafavi et al., 2008) to predict the gene-gene interaction network (Warde-Farley et al., 2010).

2.7. Protein structural and mutational analysis

For the prediction of structural and mutational effects in the targeted gene, we utilized the HOPE server to analyze SNP consequences, structural confirmations and function of the target protein (Venselaar et al., 2010). The protein mutants were modelled using the SWISS-MODEL server (Guex and Peitsch, 1997). To predict the DEFB1 protein-binding residues linked to the particular nsSNPs, the SCRIBER webserver was applied to analyze the wild-type protein sequence by quantifying putative propensities to bind proteins (Zhang and Kurgan, 2019). The protein-binding residue prediction was followed by the additional analysis to determine the ligand 3D binding sites in DEFB1. This was carried out by the Pocket Picker algorithm with shape descriptors as the PyMol plugin (Weisel et al., 2007).

2.8. Molecular docking and dynamics simulation

The plasma membrane lipid phosphatidylinositol 4,5-bisphosphate (PIP2) was obtained from the Protein Data Bank as a crystal structure of human DEFB2 in complex with PIP2 (Jarva et al., 2018). The AutoDock molecular docking algorithm and the Rosetta protein-ligand docking protocol were implemented using the PIP2 binding site to calculate binding affinities (ΔG_{bind} and E^{ros}). The receptor and ligand structure preparations for molecular docking included Gasteiger partial charges assignment and rotatable bonds definition according to the standard protocol published elsewhere (Shityakov et al., 2014). AutoDock v.4.2.5.1 was used in the study since its previous version incorrectly calculates part of the intermolecular desolvation energy term (Goodsell et al., 1996). The docking grid with a dimension size of 60 × 60 × 60 Å and a grid spacing of 0.375 Å were used in the study. All molecular dynamics (MD) simulations were performed using the AMBER 16 package with the FF99SB and GAFF force fields for DEFB1

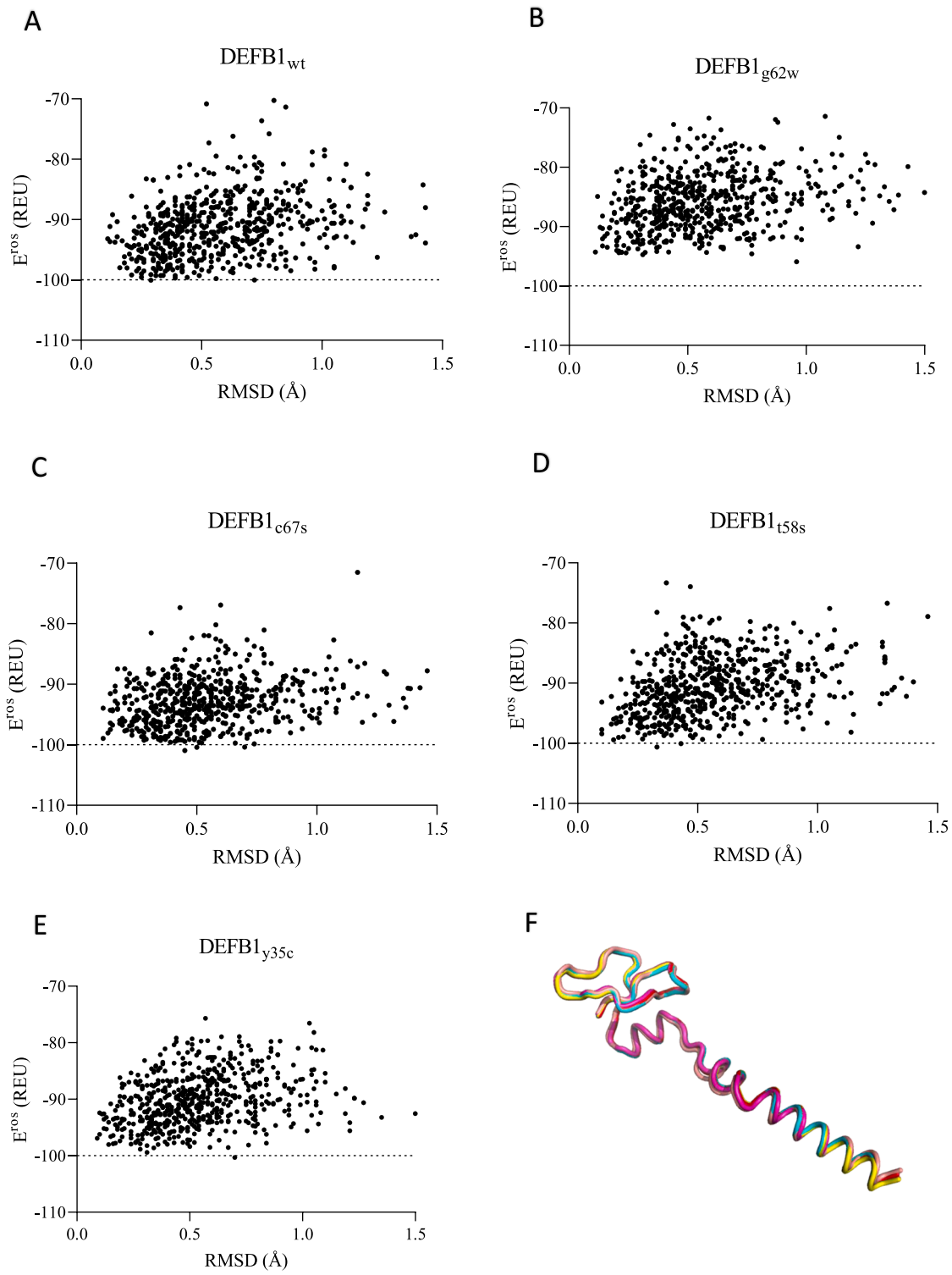

Fig. 1. All-atom RMSD vs. Rosetta energy (E^{ros}) scatter plots for the analysed protein molecules including wild-type (wt) and mutants of DEFB1 (A-E). 3D structural alignment (F) of the mutant molecules upon the reference structure (wt) is represented as a tube model. The reference DEFB1 structure is coloured in red. REU is abbreviated as Rosetta energy units. The energy threshold ($E^{ros} = -100$ REU) is depicted as a dashed line.

and PIP2 (Case et al., 2005). The Antechamber module of AmberTools was employed to calculate the Gasteiger partial charges of PIP2. The systems were solvated with the TIP3P water models and neutralized by adding the Na+ ions using the tLEap input script available from the AmberTools package. Long-range electrostatic interactions were modeled via the particle-mesh Ewald method (Essmann et al., 1995). The SHAKE algorithm (Miyamoto and Kollman, 1992) was applied to constrain the length of covalent bonds, including the hydrogen atoms. Langevin thermostat was implemented to equilibrate the temperature of the system at 310 K. A 2.0-fs time step was used in all of the MD setups. For the minimization and equilibration (NVT and NPT ensembles) phases, 100,000 steps and a 1-ns period were used, respectively. Finally, 100-ns classical MD simulations, with no constraints as NPT ensemble, were performed for each of the protein -ligand complexes. Additionally,

Table 3
conserved Summary of conserved areas, SNP effects and post-translational modifications.

Mutation	Conserved areas	SNP effect	PTM
C67S	region 9-highly conserved	Functional	Disulfide-linkage
T58S	region 8-highly conserved	Functional	Amidation, O-linked mechanism of glycosylation
G62W	region 6-intermediate conserved	None	None
Y35C	region 7-highly conserved	None	Amidation mechanism

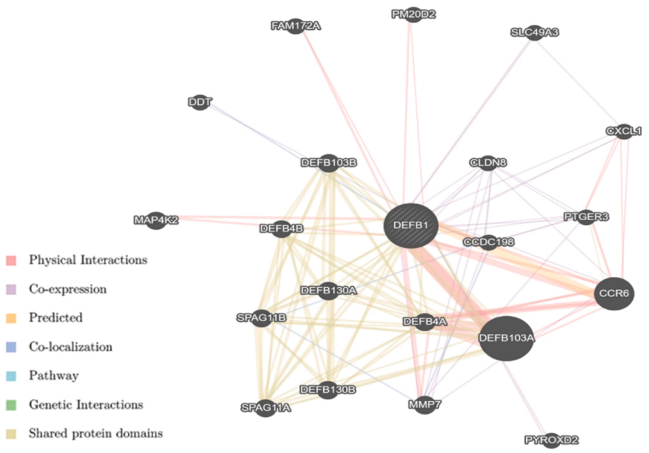

Fig. 2. Gene association network of human DEFB1 (Association with other genes: Physical interactions: 77.84%, Co-expression: 8.01%, Predicted: 5.37%, Co-localization: 3.83%, Genetic Interactions: 2.87%, Pathways: 1.88%, Shared protein domains: 0.60%).

the MD protocol was combined with the Poisson–Boltzmann (MM-PBSA) or generalized Born (MM-GBSA) molecular mechanics approach augmented with the hydrophobic solvent-accessible surface area term (Shityakov et al., 2021).

3. Results

3.1. Data mining

In this study, the total SNP number in the human DEFB1 gene was detected to be 4024. Among them, there were 70 missense, 22 synonymous, and 35 other SNPs in the 5'-UTR region o, and 40 SNPs in the 3'-UTR region There were 2650 SNPs of upstream, downstream and intronic regions of transcripts.

3.2. Identification of deleterious nsSNPs in human DEFB1

The 70 missense SNPs were further examined through the SIFT tool; and 12 ns-SNPs deleterious were identified with the tolerant values of 0.05. Next, the PolyPhenV2 tool with a default tolerant accuracy scores provided five SNPs. The other algorithms, including Provean Tool, SNAP, PHD-SNP, SNP&GO revealed only four deleterous SNPs related to DEFB1 as shown in Table 1.

3.3. Protein stability prediction

According to the data obtained from the I-Mutant tool and Rosetta protocol, a very slight stability increase was detected for all the mutated proteins except for the G62W mutant predicted with highest ΔΔG and positive ΔE^{ros} (Table 2; Fig. 1[A-E]). Moreover, the RMSD values were also determined to become highest for the less stable protein structure (Fig. 1[F]). These results indicated no significant impact of the analyzed mutations on the protein stability except for G62W.

Table 4
DEFB1 structural and mutational analysis.

Mutation	Interaction	Location	SNP effect
C67S	There is a difference in hydrophobicity between the wild-type and mutant residue. Any hydrophobic interactions within the protein or on its surface will be destroyed as a result of the mutational process.	Buried, and core organization of its domains.	The variants residue is conserved position and likely as deleterious.
T58S	There is a size difference between wild-type and mutant amino acids. Smaller mutations may result in a lack of interaction.	The amino acid variation in the surface area of specific domains with undefined function.	The polymorphism is not damaging.
G62W	There is a size difference between wild-type and mutant amino acids. There is a greater amount of mutant residue, which might cause bumps. This residue's torsion angles are very steep. These torsion angles are only possible with glycine, since any other residue are push the back-bone into an erroneous vindication, causing local structural disruption.	Unknown function of a mutated amino acid in protein surface domains.	The residue of amino acid is conserved.
Y35C2	The size of wild-type and mutant amino acids differs. Because the mutant residue is smaller, interactions may be lost. The hydrophobicity of the mutant and wild-type residues is different. At this location, the mutation introduces a more hydrophobic residue. This can cause hydrogen bonding to break down and/or interfere with proper folding.	Unknown function of a variant amino acid on the surface domain.	Variant is found without being deleterious to protein.

3.4. Analysis of nsSNP conservation

Conserved amino acids during evolution are important for the correct function of a protein but some deleterious mutations are more frequent in conserved sequences. Selected nsSNPs were predicted to be in the highly conserved regions 6–9 having structural and functional effects on the DEFB1 protein (Table 3).

3.5. Post-translational modification analysis

It is well known that PTMs play a vital role in protein folding and disintegration, transcriptional regulation, and a variety of biochemical processes (Ramazi and Zahiri, 2021). To identify the PTM sites in the wt and mutated proteins, we implemented the ModPred algorithm as a unified sequence-based predictor. Additionally, the information on

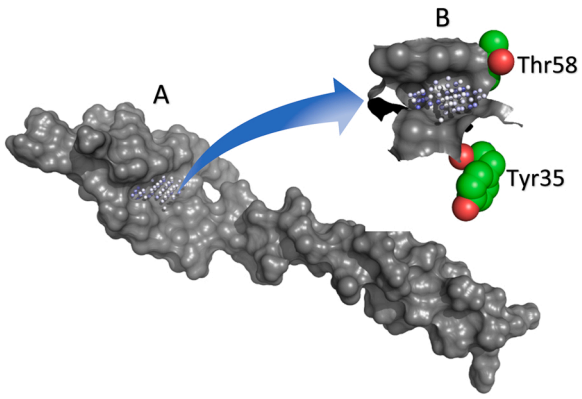

Fig. 3. The PIP2 binding site in the wild-type DEFB1 protein is predicted with the PocketPicker algorithm and depicted in dots (A). The mutated amino acid residues associated with the binding site are shown in spheres and coloured according to their atomic composition (B). The molecular surface is coloured in grey.

gene-to-gene functionality and interaction with DEFB1 was obtained from the GeneMANIA server as shown in Fig. 2. The gene-gene network analysis for DEFB1 revealed two additional hub gene nodes as DEFB103A and CCR6 (C-C Motif Chemokine Receptor 6) connected mainly due to their physical interaction (77.84%). The latter gene is responsible for the metastatic spread of gastrointestinal malignancies, where the up-regulated CCR6 expression could be a hallmark of colorectal cancer (Frick et al., 2016; Rubie et al., 2014).

3.6. Protein structural and mutational analysis

To have some insight into the structural effects of the predicted mutations in DEFB1, we utilized the HOPE algorithm as a fully automatic program that analyzes the structural and functional effects of point mutations. In addition, the SCRIBER program was able to assess the protein (wt and mutated variants) binding to DNA, RNA, ligand, or other proteins (Supplementary material). As a result, most of the mutated variants represented some differences in size between wt and mutant amino acids with their unknown function on the protein domains (Table 4). Further structural analysis of protein-ligand binding sites indicated that only the Tyr35 and Thr58 residues were identified as binding site-forming residues. Therefore, their possible mutations (Y35C and T58S) could lead to an impairment of protein-ligand interaction. This analysis was also performed to identify surface binding pockets and

Table 5
Summary of binding affinities and energetic terms obtained from analysis of of molecular docking results or 100 ns MD trajectories of wt or mutated DEFB1 proteins bound to PIP2). Energy values are measured in kcal/mol ± SD for ΔG and REU (Rosetta Energy Units) for E^{ros}.

Complex	ΔG_{bind}	ΔG_{PBSA}	ΔG_{GBSA}	E^{ros}
wt-PIP2	− 7.07	− 31.09 ± 5.28	− 40.7 ± 4.05	− 122.68
Y35C-PIP2	− 6.07	− 9.05 ± 2.86	− 25.51 ± 5.3	− 118.06
T58S-PIP2	− 6.04	− 9.95 ± 3.46	− 14.19 ± 2.56	− 120.01

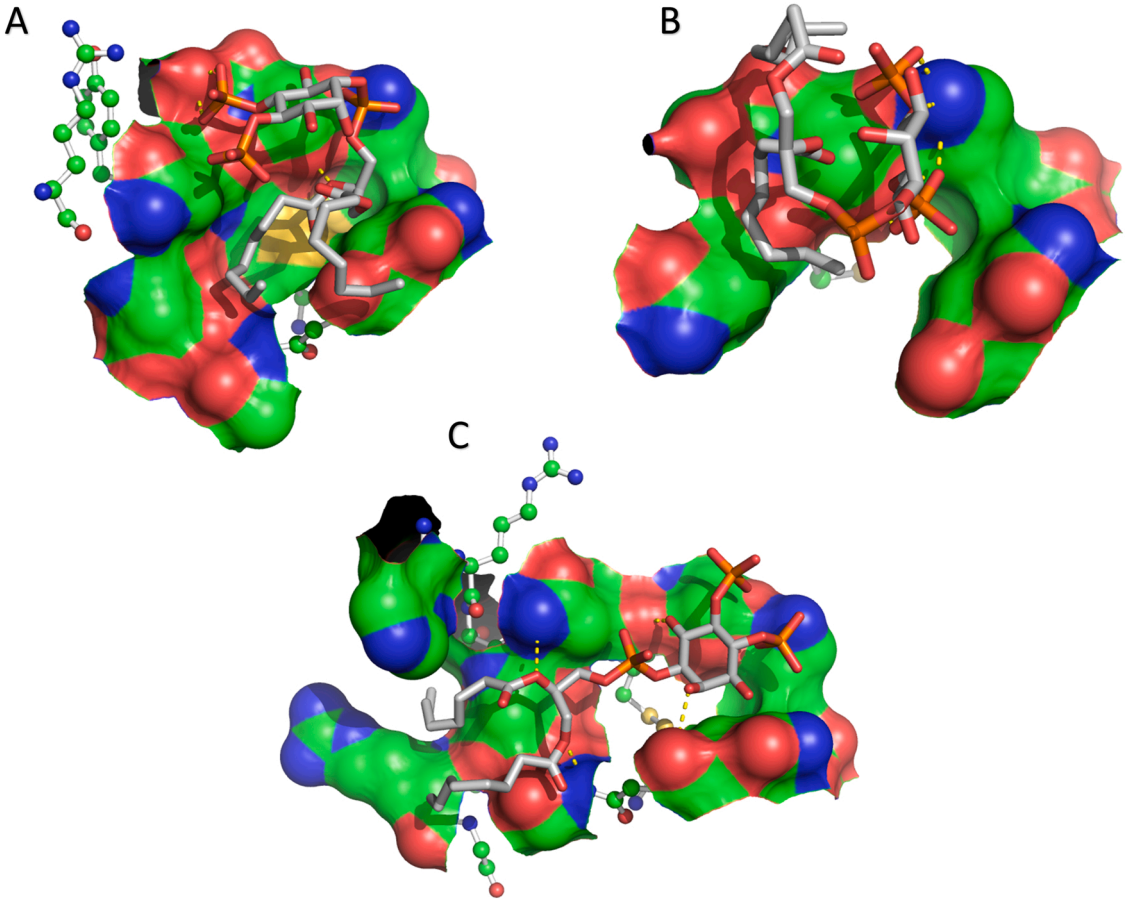

Fig. 4. 3D binding modes predicted with the Autodock algorithm for the wt (A), Y35C (B), and T58S (C) proteins bound to the bacterial PIP2 molecule. The protein-ligand binding site is shown by the molecular surface and coloured according to the protein atomic composition. The protein residues are drawn as ball-and-stick models. Hydrogen bonds are visualized as a dashed line. The ligand molecules are depicted in sticks partially coloured in grey, and hydrogen atoms are removed to enhance clarity.

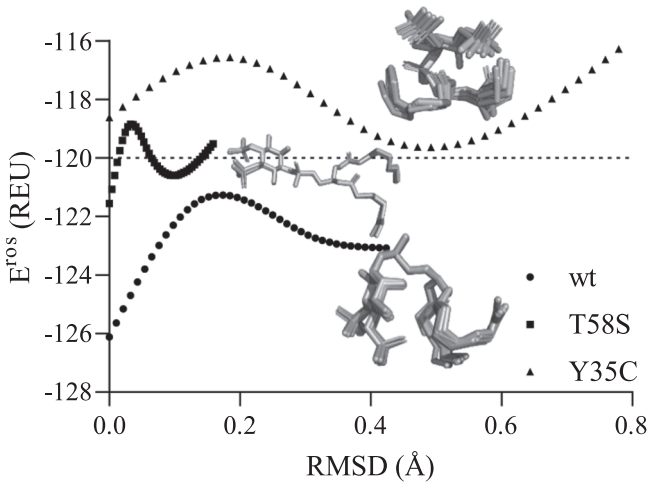

Fig. 5. Binding modes predicted with the Monte Carlo algorithm (Rosetta protein-ligand docking protocol) for the wt, Y35C, and T58S proteins bound to PIP2. The energy threshold ($E^{ros} = -120$ REU) is depicted as dashed line. The aligned ligand conformations are shown as stick models.

occluded cavities by the PocketPicker algorithm, where the same residues were found to be associated with the binding site geometry (Fig. 3 [A, B]).

3.7. Molecular docking and dynamics simulations

It was previously discovered that defensin molecules could possess a broad spectrum of antimicrobial activities via targeting the PIP2 molecule. to permeabilize bacterial membranes (Jarva et al., 2018). Therefore, rigid-flexible molecular docking was performed to evaluate the PIP2 binding to the human wt DEFB1 structure and its Y35C, and T58S mutants (Fig. 4[A–C]). As a result, the wt protein was found to have the best affinity to PIPI2 with the binding energy (ΔG_{bind}) value of − 7.07 kcal/mol (Table 5), which was higher for the mutant forms (−6.07 and −6.04 kcal/mol). However, the PIP2 molecule was predicted to be a binder for all analyzed proteins due to the energy values lower than − 6.0 kcal/mol threshold (Table 5), previously determined as a reference value (Shityakov and Förster, 2014). Similar results were predicted using the Monte Carlo methodology integrated in the Rosetta protein-ligand docking protocol, showing the highest affinity ($E^{ros} = -122.68$ REU) for the wt protein molecule to bind PIP2 (Fig. 5, Table 5). To validate and confirm molecular docking results, the free energy of binding based on implicit solvation models (ΔG_{PBSA} and ΔG_{GBSA}) was calculated for the DEFB1-PIP2 complexes. The MM-PBSA/GBSA calculations (Table 5), using 100 ns MD trajectories, confirmed the previous data completely (GBSA) and partially (PBSA), revealing much higher binding affinities for the wt DEFB1-PIP2 complex.

To investigate the movements of the studied complexes, the root-mean-square deviation (RMSD) and fluctuation (RMSF) values together with the radius of gyration (Rg) concerning the initial conformation were plotted versus time (Fig. 6[A–E]). The receptor RMSD values were relatively high (10–15 Å) stabilized after about 25 ns (Fig. 6 [A]). The ligand RMSD values were determined to be in the range from 2 to 4 Å being stabilized almost instantly for Y35C and after 25 ns for wt and T58S (Fig. 6[B]). The receptor RMSF values produced various picks associated with the high flexibility of the DEFB1 residues. (Fig. 6[C]). The atomic fluctuations of ligand showed similar pattern to the receptor RMSF profiles (Fig. 6[D]). The Rg values, associated with the receptor compactness, were starting to decrease instantaneously and stabilized after 10 ns for wt and 25 ns for the mutant forms (Fig. 6[E]). Finally, the number of H-bonds between the proteins and their ligand was assessed to find the H-binding contribution to the affinity. This parameter was the highest for wt with more H-bonds formed (Fig. 6[F]), which corroborates with the previous findings.

4. Discussion

Multiple computational techniques were employed in this study to predict and evaluate the deleterious nsSNPs of the DEFB1 gene. As a result, four nsSNPs have been found to be deleterious, such as C67S, T58S, G62W, and Y35C. The stability analysis revealed the slight increase in the protein stability of the mutant forms except for the G62W mutant. In reality, the nsSNPs mechanisms that affect protein stability are very complex depending on geometrical constraints, various physicochemical effects, and a formation or disruption of salt bridges, disulfide, and hydrogen bonds (Teng et al., 2009). Moreover, disulfide bridges are usually necessary for the antimicrobial activity of antimicrobial peptides, such as defensins (Lei et al., 2019). In our case, the ModPred algorithm had revealed the nsSNPs impact on PTM via amidation-glycosylation mechanisms and disulfide linkages. Indeed, if the PTMs are affected by nsSNPs, it would change signaling pathways, such as phosphorylation, ubiquitination, and acetylation leading to a specific pathology (Kim et al., 2015). On the other hand, the gene-gene interactions are also important to consider as they also influence innate immunity (Basu et al., 2012). The GeneMANIA algorithm had predicted that the DEFB1 gene could interact with CCR6, which encodes a C-C motif chemokine receptor 6 produced by dendritic cells and memory T-cells (Liao et al., 1999). As the DEFB1 molecule is chemoattractive for immature dendritic cells, it could induce the CCR6 chemokine receptor activation (Andresen et al., 2011).

It is well known that the DEFB1 down-regulation is mainly associated with different cancer pathologies, such as conventional clear cell carcinoma (Donald et al., 2003). In this study, 90% of cancer cells have shown the cancer-specific loss of DEFB1 protein (Donald et al., 2003). On the contrary, the loss-of-function mutation in the DEFB1 gene demonstrated no obvious defect in the innate immunity of mice due to the redundancy of defensin molecules (Morrison et al., 2002). However, this phenomenon has never been determined in humans.

A number of theoretical and computer simulation approaches have been established to describe defensin interaction with lipid membranes (de Paula and Valente, 2018). In some computational studies, the interaction of human defensin type 3 (hBD-3) with different PIP2-containing membranes has been investigated (Zhang, 2021). It was found that the hBD-3 analog binds on the mixed lipid membrane in the two-loop regions (Zhang, 2021). Lee and co-authors examined the adsorption mechanism of hBD3 using Gram-negative (GN), Gram-positive (GP), and mammalian membrane models, via MD simulations (Lee et al., 2016). They concluded that Arg17, Arg36, and Arg38 as the key residues could form both polar and nonpolar interactions to exhibit antibacterial activity. Moreover, a high concentration of POPG lipids in the GP membrane leads to strong electrostatic interaction with hBD-3 stabilizing its adsorption (Lee et al., 2016). On the other hand, site-directed mutant porcine β-beta defensin type 2 proteins showed different antimicrobial activities against *E. coli* and *S. aureus* evaluated by the turbidimetric method (Huang et al., 2015). This antibacterial effect could be attributed to the different distribution of positive charges in the protein molecule or most likely to a varying geometry of the protein-ligand binding sites in the protein molecule (Huang et al., 2015). Thus, the mutation impact on genes encoding innate immunity factors is important to understand the molecular mechanisms of antibacterial action, leading to the strategic design of new pharmacologically active compounds.

5. Conclusion

SNPs can be used as important genetic biomarkers, such as the DEFB1-associated nsSNPs. They could also help to evaluate protein function and their possible contribution to various pathological

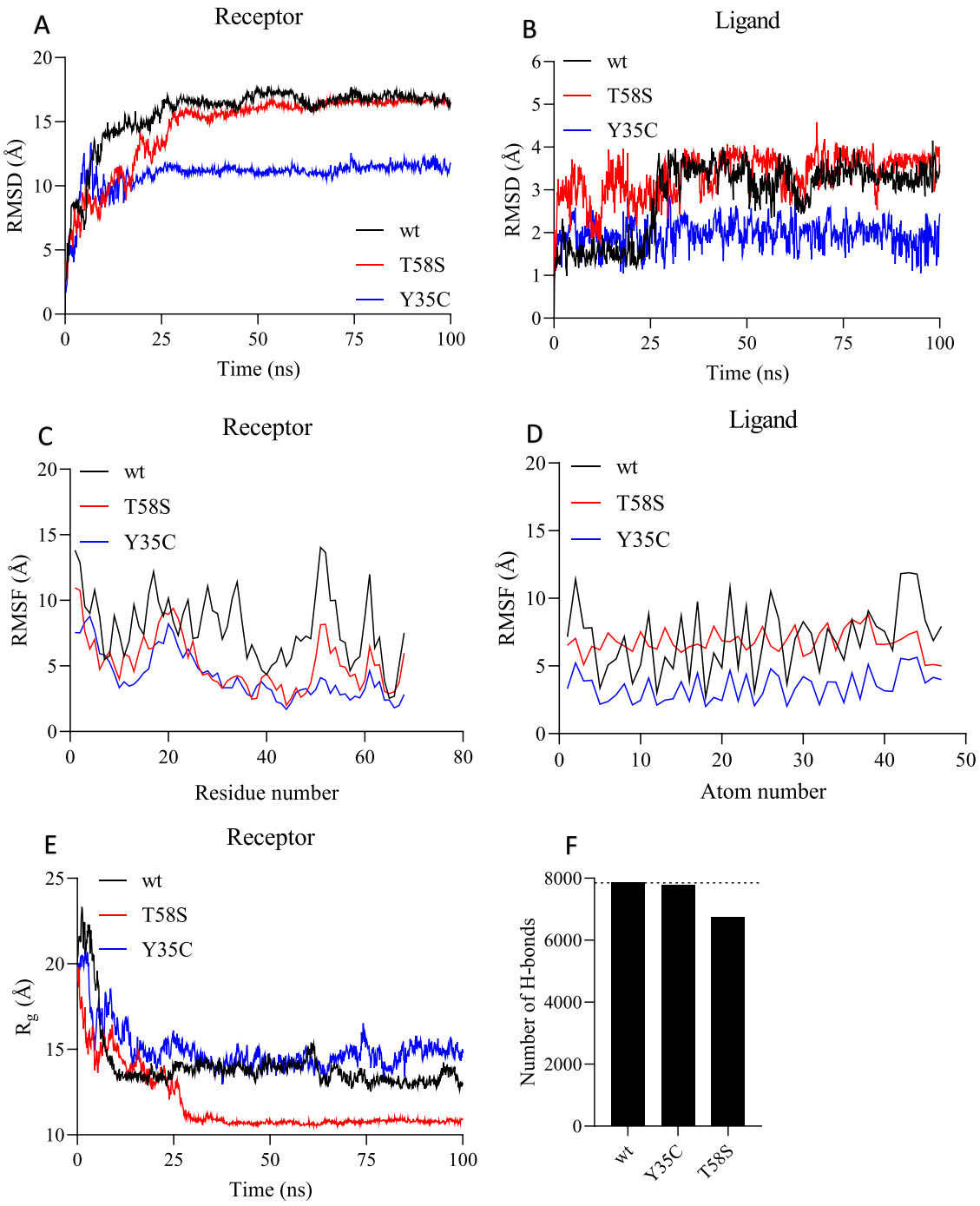

Fig. 6. RMSD (A, B), RMSF (C, D), Rg (E), and number of H-binds (F) involved in H-bonding calculated using the wt and mutated DEFB1-PIP2 complexes during 100 ns MD simulations.

processes. These biomolecules are known to be a part of innate immunity displaying antimicrobial and immunomodulatory activities and functioning as little cationic antimicrobial peptides. In this study, we implement cutting-edge computational algorithms to identify deleterious SNPs in the DEFB1 gene that could potentially change the 3D geometry of protein-ligand binding sites. As a result, 4 most important nsSNPs in the DEFB1 gene were predicted (C67S, T58S, G62W, and Y35C) and only two of them were discovered to be linked to the binding site-forming residues, such as Thr58 and Tyr35. In addition, molecular docking and molecular dynamics experiments demonstrated the decrease of DEFB1 binding affinity to PIP2 due to these mutations. Altogether, this computational study of nsSNPs in DEFB1 provides more understanding of how these missense mutations could impair or change protein functions by altering the DEFB1-PIP2 binding site.

CRediT authorship contribution statement

MMF and SS conceived the idea and design of the review. SU, SA, and TAJ refined and expanded the design. SU and MMF wrote the manuscript with input from all authors. All authors read and approved the final manuscript.

Data Availability

The data that support the findings of this study are available upon request from the authors.

Acknowledgments

The work was carried out with the support of the Ministry of Science and Higher Education of the Russian Federation (agreement no. 075-15-2021-1349). Special thanks are extended to Anastasia Nenashkina from the Infochemistry Scientific Center, ITMO University for her assistance in designing of the graphical abstract, and to the School of Science and Engineering, Department of Computer Science, Università degli studi di Verona, Verona, Italy.

Conflict of interest

The authors declare no competing interests.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.compbiolchem.2022.107669.

References

Abbas, R.A.A., et al., 2019. Computational analysis of single nucleotide polymorphism (SNPs) in HumanSLC5A1 Gene. Int. J. Biomed. Sci. Eng. 7 (4), 85.
Adzhubei, I., Jordan, D.M., Sunyaev, S.R., 2013. Predicting functional effect of human missense mutations using PolyPhen-2. Curr. Protoc. Hum. Genet. 76 (1), 7.20, 1-7.20. 41.
Alwi, Z.B., 2005. The use of SNPs in pharmacogenomics studies. Malays. J. Med. Sci.: MJMS 12 (2), 4–12.
Andresen, E., et al., 2011. Increased expression of beta-defensin 1 (DEFB1) in chronic obstructive pulmonary disease. PLoS One 6 (7).
Basu, M., et al., 2012. Gene-gene interaction and functional impact of polymorphisms on innate immune genes in controlling plasmodium falciparum blood infection level. PLoS One 7 (10).
Case, D.A., et al., 2005. The amber biomolecular simulation programs. J. Comput. Chem. 26 (16), 1668–1688.
Choi, Y., Chan, A.P., 2015. PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. Bioinformatics 31 (16), 2745–2747.
Consortium, U., 2015. UniProt: a hub for protein information. Nucleic Acids Res. 43 (D1), D204–D212.
Donald, C.D., et al., 2003. Cancer-specific loss of beta-defensin 1 in renal and prostatic carcinomas. Lab. Investig. 83 (4), 501–505.
Essmann, U., et al., 1995. A smooth particle mesh ewald method. J. Chem. Phys. 103 (19), 8577–8593.
Ghosh, P., et al., 2021. In-silico analysis of non-synonymous SNPs of human LDLR gene and their impact on familial hypercholesterolemia. Gene Rep. 23, 101127.
Glaser, F., et al., 2003. ConSurf: identification of functional regions in proteins by surface-mapping of phylogenetic information. Bioinformatics 19 (1), 163–164.
Goodsell, D.S., Morris, G.M., Olson, A.J., 1996. Automated docking of flexible ligands: applications of AutoDock. J. Mol. Recognit. 9 (1), 1–5.
Guex, N., Peitsch, M.C., 1997. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis 18 (15), 2714–2723.
Heimlich, D.R., Harrison, A., Mason, K.M., 2014. Host antimicrobial peptides in bacterial homeostasis and pathogenesis of disease. Antibiotics 3 (4), 645–676.
Herraez, D.L., et al., 2009. Genetic variation and recent positive selection in worldwide human populations: evidence from Nearly 1 Million SNPs. PLoS One 4 (11).
Huang, X.X., et al., 2015. Antimicrobial characterization of site-directed mutagenesis of porcine beta defensin 2. PLoS One 10 (2).
Jarva, M., et al., 2018. Human beta-defensin 2 kills Candida albicans through phosphatidylinositol 4,5-bisphosphate-mediated membrane permeabilization. Sci. Adv. 4 (7).
Jumper, J., et al., 2021. Highly accurate protein structure prediction with AlphaFold. Nature 596 (7873), 583 (-+).
Jurevic, R.J., et al., 2003. Single-nucleotide Polymorphisms (SNPs) in human beta-defensin 1: High-throughput SNP assays and association with Candida carriage in type I diabetics and nondiabetic controls. J. Clin. Microbiol. 41 (1), 90–96.
Kim, Y., et al., 2015. Detection and analysis of disease-associated single nucleotide polymorphism influencing post-translational modification. Bmc Med. Genom. 8.
Kumar, P., Henikoff, S., Ng, P.C., 2009. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat. Protoc. 4 (7), 1073–1081.
Lee, J., Jung, S.W., Cho, A.E., 2016. Molecular Insights into the Adsorption Mechanism of Human beta-Defensin-3 on Bacterial Membranes. Langmuir 32 (7), 1782–1790.
Lehrer, R.I., Bevins, C.L., Ganz, T., 2005. Defensins and other antimicrobial peptides and proteins. Mucosal Immunol. 95–110.
Lei, J., et al., 2019. The antimicrobial peptides and their potential clinical applications. Am. J. Transl. Res. 11 (7), 3919–3931.
Liao, F., et al., 1999. CC-chemokine receptor 6 is expressed on diverse memory subsets of T cells and determines responsiveness to macrophage inflammatory protein 3 alpha. J. Immunol. 162 (1), 186–194.
Machado, L.R., Ottolini, B., 2015. An evolutionary history of defensins: a role for copy number variation in maximizing host innate and adaptive immune responses. Front. Immunol. 6.
Meade, K.G., O'Farrelly, C., 2019. β-Defensins: farming the microbiome for homeostasis and health. Front. Immunol. 9 (3072).
Miyamoto, S., Kollman, P.A., 1992. Settle - an analytical version of the shake and rattle algorithm for rigid water models. J. Comput. Chem. 13 (8), 952–962.
Morrison, G., et al., 2002. Characterization of the mouse beta defensin 1, Defb1, mutant mouse model. Infect. Immun. 70 (6), 3053–3060.
Mostafavi, S., et al., 2008. GeneMANIA: a real-time multiple association network integration algorithm for predicting gene function. Genome Biol. 9 (1), 1–15.
de Paula, V.S., Valente, A.P., 2018. A dynamic overview of antimicrobial peptides and their complexes. Molecules 23 (8).
Pejaver, V., et al., 2014. The structural and functional signatures of proteins that undergo multiple events of post-translational modification. Protein Sci. 23 (8), 1077–1093.
Prado-Montes de Oca, E., 2010. Human β-defensin 1: a restless warrior against allergies, infections and cancer. Int. J. Biochem. Cell Biol. 42 (6), 800–804.
Ramazi, S., Zahiri, J., 2021. Post-translational modifications in proteins: resources, tools and prediction methods. Database 2021.
Ramelot, T.A., et al., 2009. Improving NMR protein structure quality by Rosetta refinement: a molecular replacement study. Proteins-Struct. Funct. Bioinforma. 75 (1), 147–167.
Shastry, B.S., 2009. SNPs: Impact on Gene Function and Phenotype, in Single Nucleotide Polymorphisms: Methods and Protocols, Second Edition, A.A. Komar, Editor, p. 3–22.
Sherry, S.T., et al., 2001. dbSNP: the NCBI database of genetic variation. Nucleic Acids Res. 29 (1), 308–311.
Shityakov, S., et al., 2014. Ionization states, cellular toxicity and molecular modeling studies of midazolam complexed with trimethyl-beta-cyclodextrin. Molecules 19 (10), 16861–16876.
Shityakov, S., et al., 2020. Novel approach for characterizing propofol binding affinities to serum albumins from different species. Acs Omega 5 (40), 25543–25551.
Shityakov, S., et al., 2021. Scaffold searching of FDA and EMA-approved drugs identifies lead candidates for drug repurposing in alzheimer's disease. Front. Chem. 9.
Shityakov, S., Förster, C., 2014. In silico predictive model to determine vector-mediated transport properties for the blood-brain barrier choline transporter. Adv. Appl. Bioinforma. Chem.: AABC 7, 23–36.
Sim, N.-L., et al., 2012. SIFT web server: predicting effects of amino acid substitutions on proteins. Nucleic Acids Res. 40 (W1), W452–W457.
Subbiah, H.V., Babu, P.R., Subbiah, U., 2020. In silico analysis of non-synonymous single nucleotide polymorphisms of human DEFB1 gene. Egypt. J. Med. Hum. Genet. 21 (1).
Teng, S.L., et al., 2009. Modeling effects of human single nucleotide polymorphisms on protein-protein interactions. Biophys. J. 96 (6), 2178–2188.
Venselaar, H., et al., 2010. Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces. BMC Bioinforma. 11 (1), 1–10.
Warde-Farley, D., et al., 2010. The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. Nucleic Acids Res. 38 (suppl_2), W214–W220.
Weisel, M., Proschak, E., Schneider, G., 2007. PocketPicker: analysis of ligand binding-sites with shape descriptors. Chem. Cent. J. 1.
Zhang, J., Kurgan, L., 2019. SCRIBER: accurate and partner type-specific prediction of protein-binding residues from proteins sequences. Bioinformatics 35 (14), I343–I353.
Zhang, L.Q., 2021. Interaction of human beta defensin type 3 (hBD-3) with different PIP2-containing membranes, a molecular dynamics simulation study. J. Chem. Inf. Model. 61 (9), 4670–4686.
Frick, V.O., Rubie, C., Keilholz, U., Ghadjar, P., 2016. "Chemokine/chemokine receptor pair CCL20/CCR6 in human colorectal malignancy: An overview". World Journal of Gastroenterology 22 (2), 833–841.
Rubie, C., Kruse, B., Frick, V.O., Kölsch, K., Ghadjar, P., Wagner, M., Grässer, F., Wagenpfeil, S., 2014. "Chemokine receptor CCR6 expression is regulated by miR-518a-5p in colorectal cancer cells". Journal of Translational Medicine 12–48.