Cell. Mol. Biol. (ISSN: 1165-158X)

Journal Homepage: www.cellmolbiol.org

Cellular and Molecular Biology

Original Article

Structural dynamics of PCSK9 loss-of-function variants: implications for LDL cholesterol regulation and cardiovascular risk

Mohammed Tarique[1], Ayyub Ali Patel[2], Ayaz Khurram Mallick[2], Rania A. Hussien[3], Tarek Mahmoud Mirdad[4], Zia-Ul-Sabah[5], Rasha Mirdad[6], Imad Mohammad Alghawanmeh[2], Mohamed Babiker Abdelrouf Awadelkarim[2], Mohammad Azhar Kamal[7*], Hani Alothaid[8]

[1] Almanac Life Science India Private Limited, Jamia Nagar, New Delhi-110025, India.
[2] Department of Clinical Biochemistry, College of Medicine, King Khalid University, Abha, Saudi Arabia
[3] Department of Chemistry, Faculty of Science, Al-Baha University, Al-Baha, Kingdom of Saudi Arabia
[4] Department of Orthopedics, College of Medicine, King Khalid University, Abha, Saudi Arabia
[5] Department of Medicine, College of Medicine, King Khalid University, Abha, Saudi Arabia
[6] Department of Surgery, College of Medicine, King Khalid University, Abha, Saudi Arabia
[7] Department of Pharmaceutics, College of Pharmacy, Prince Sattam Bin Abdulaziz University, Al-Kharj 11942, Saudi Arabia
[8] Department of Basic Medical Science, Faculty of Applied Medical Sciences, Al-Baha University, Al-Baha Province, Saudi Arabia

Article Info

Article history:

Received: September 08, 2025
Accepted: November 29, 2025
Published: December 31, 2025

Use your device to scan and read the article online

Abstract

Cardiovascular disease (CVD) is a leading global cause of mortality, and understanding its underlying mechanisms is crucial for developing effective interventions. The liver-derived protein PCSK9 (proprotein convertase subtilisin/kexin-type-9) plays a vital role in regulating lipoprotein metabolism by binding to the low-density lipoprotein receptor (LDLR) and promoting its lysosomal degradation, ultimately reducing low-density lipoprotein (LDL) clearance. Loss-of-function (LOF) variants in PCSK9 are associated with decreased LDL cholesterol (LDL-C) levels, suggesting that these variants may contribute to a lower risk of cardiovascular events. Our computational analysis of PCSK9 LOF variants revealed significant alterations in stability, flexibility, and free energy compared to the native protein. Protein-protein docking studies of both wildtype and mutant PCSK9 with LDLR demonstrated variations in binding energy and interacting residues. Notably, while the binding cavity remained the same as that of the wildtype, all variants exhibited distinct binding interactions. Molecular dynamics simulations further highlighted increased flexibility and solvent exposure in the mutant protein complexes. These findings indicate that LOF variants in PCSK9 induce substantial structural changes, leading to a decreased affinity for LDLR binding.

Keywords: Cardiovascular disease, low-density lipoprotein cholesterol, mutation, protein-protein docking, molecular dynamics simulation

1. Introduction

Cardiovascular disease (CVD) remains a global health challenge, claiming countless lives annually. By the year 2030, approximately 23.6 million individuals will lose their lives annually due to CVD. CVDs encompass a diverse group of medical conditions affecting the heart and circulatory system and contributing to the considerable burden of CVD-related morbidity and mortality. CVDs are characterized as chronic ailments that progress gradually over an individual's lifetime and frequently remain asymptomatic for extended periods. Typically, symptoms only manifest in advanced stages, and in some cases, the initial indication may unfortunately be sudden and severe [1].

Cholesterol metabolism is a complex process essential for maintaining cellular functions and overall lipid homeostasis, involving the synthesis, transport, and regulation of cholesterol within the body. Low-density lipoprotein (LDL) cholesterol, often referred to as "bad cholesterol," is primarily transported in the bloodstream by LDL particles, which are cleared by the liver through receptor-mediated endocytosis via the low-density lipoprotein receptor (LDLR). Proprotein convertase subtilisin/kexin type 9 (PCSK9) plays a crucial role in this regulatory mechanism by binding to LDLR and promoting its degradation in lysosomes, thereby reducing the number of LDLR available on the cell surface. This PCSK9-mediated degradation limits the liver's ability to clear LDL from circulation, resulting in increased serum LDL cholesterol levels, which is a significant risk factor for cardiovascular diseases. Understanding the interplay between PCSK9 and LDLR is

* Corresponding author.
E-mail address: ma.kamal@psau.edu.sa (M. A. Kamal).
Doi: http://dx.doi.org/10.14715/cmb/2025.71.12.15

vital for developing therapeutic strategies aimed at lowering LDL cholesterol levels and reducing the risk of atherosclerosis and heart disease.

Among the intricate landscape of factors influencing CVD genetics is one of the major contributing factors. Among many genetic contributors, PCSK9 (proprotein convertase subtilisin/kexin-type-9), a liver-derived protein involved in the degradation of LDLR (low-density lipoprotein receptors) in the liver, has emerged as a key player [2]. The protein assumes a pivotal role in the finely tuned orchestration of lipoprotein metabolism, orchestrating a cascade of events that profoundly affect cholesterol levels within the body. Specifically, PCSK9 interacts with LDLR and exerts its influence by promoting the lysosomal degradation of PCSK9 in the liver [3, 4]. This degradation process, in turn, curtails the clearance of low-density lipoprotein cholesterol (LDL-C), informally known as "bad cholesterol," from the bloodstream.

The presence of cholesterol in the bloodstream is influenced to certain extent by individuals dietary choices. However, the dominant factor governing cholesterol levels is genetics. Certain variations within the PCSK9 gene are associated with increased PCSK9 protein production, resulting in elevated cholesterol levels. Specific gain-of-function (GOF) mutations within this gene can lead to exceptionally high levels of LDL cholesterol, a condition recognized as familial hypercholesterolemia. Conversely, there are other PCSK9 variants that cause only a modest increase in LDL cholesterol levels, elevating them slightly above the normal level [5].

Researches across the globe have demonstrated that certain variations in PCSK9 (referred to as loss-of-function (LOF) variants) are correlated with reduced cholesterol levels [6]. These genetic variants exhibit diminished functionality and trigger an increase in the number of LDL receptors within the liver. Consequently, this heightened receptor activity results in a more efficient removal of LDL particles, including cholesterol, from the bloodstream. Consequently, individuals harbouring these LOF variants tend to maintain lower LDL cholesterol levels over their lifetime, which, in turn, is associated with a reduced susceptibility to CVD [7]. Investigation of these LOF variants can provide answers to the reduction of LDL-C levels mediated by these variants in relation to CVD.

Recent findings suggest that PCSK9 not only regulates LDLR but also interacts with the metabolism of very-low-density lipoproteins (VLDL) and high-density lipoproteins (HDL). The effects of different groups of loss-of-function (LoF) variants in PCSK9, those with no activity versus those with some residual activity, can significantly influence these lipid pathways. LoF variants resulting in complete loss of PCSK9 activity can lead to a marked increase in the availability of LDLR on the liver surface. This heightened LDLR expression not only enhances the clearance of LDL particles but can also affect VLDL metabolism. Increased uptake of VLDL may result in lower circulating triglyceride levels and potentially influence HDL levels as well. This reduction in VLDL can shift the balance toward a healthier lipid profile, possibly leading to lower cardiovascular risk. Conversely, LoF variants with some residual activity may have a more nuanced impact on lipid metabolism. While these variants may still increase LDLR levels and promote LDL clearance, their partial activity might not sufficiently enhance VLDL and HDL clearance. This scenario could lead to a moderate increase in LDL cholesterol levels, with potential dysregulation in VLDL and HDL metabolism. The effects could vary depending on the specific variant, but overall, some residual activity might limit the extent of beneficial lipid modulation observed with complete loss of function.

In the current study, a comprehensive computational investigation of these LOF PCSK9 variants is carried out to understand the structural implications of these variants. Various computational tools are utilized to assess the effect of these variants on the physicochemical properties of the protein. In addition, molecular modelling techniques such as protein-protein docking and molecular dynamics simulations are employed to gain a deeper understanding of the impact and dynamics of these genetic variations on the interactions between PCSK9 and Low-density lipoprotein receptor (LDLR). The insights derived from the current work will shed light on the critical residues mediating the crucial interactions, thereby differentiating the behaviour of the wildtype and LOF variants of PCSK9. This research represents a significant step towards unravelling the molecular intricacies of PCSK9's role in lipoprotein metabolism and its potential implications for cardiovascular disease prevention and management.

2. Materials and Methods

2.1. Mining and annotation of LOF variants

The LOF variants utilized in this study were sourced from diverse databases such as gnomAD, HGMD, and ClinVar [8,9]. Additionally, an extensive literature review was conducted to identify supplementary variants [10]. Ensembl Variant Effect Predictor (VEP, release 110) was employed to annotate all the variants based on the Human GRCh38 assembly [11]. To ensure a single prediction for each variant, Ensembl canonical transcripts were selected. Various variant consequences, including missense, frameshift, stop gain, and splice site alterations, were predicted by Ensembl. The variant frequency, as observed in any population (AF), was calculated from the combined 1000 Genome Project data. Genic intolerance was evaluated using the LOEUF plugin of VEP, which scores variants based on the "loss-of-function observed/expected upper bound fraction." Additionally, information regarding the protein domain of the variants was extracted from InterPro [12].

2.2. Sequence-based pathogenicity prediction of missense variants

The wildtype PCSK9 protein sequence (Accession number: Q8NBP7) was obtained from the Uniprot Database [13]. Mutant sequences were generated by replacing the wildtype amino acid with the variant amino acid. To assess the functional consequences of missense mutations, several prediction tools were employed, including SIFT (Sorting Intolerant from Tolerant) [14], SNAP2 (Screening of Nonacceptable Polymorphism 2) [15], Align-GVGD (Align-Grantham Variation Grantham Deviation) [16], MUpro, I-mutant 3.0 [17], and DDGun [18]. SIFT predicts the impact of non-synonymous mutations on protein function based on amino acid physicochemical properties and sequence homologies. SNAP2 is a neural-network-based method that uses trained classifiers to predict substitution effects on protein function. Align-GVGD combines biophysical amino acid characteristics and protein multiple

sequence alignments to categorize missense substitutions as enriched deleterious or enriched neutral. The prediction classes form a spectrum (C0, C15, C25, C35, C45, C55, C65) with C0 to be least likely and C65 to be most likely. MUpro employs support vector machine (SVM) and neural network methods to predict energy changes (ΔΔG) upon mutation. Similarly, I-Mutant 3.0, utilizing SVM, predicts protein stability changes based on free energy changes from single point mutations using sequence and structural data. DDGun employs a non-trained method to predict ΔΔG for single-site and multiple-site variations. DUET integrates smCSM and SDM through a consensus prediction approach using SVM, incorporating 3D structural information for Gibbs free energy prediction [19]. Additionally, DeepREx-WS utilizes deep learning techniques to classify residues in the sequence as buried or exposed [20].

2.3. Wildtype and mutant protein structure prediction

The wildtype protein structure was modelled using Modeller 10.4, employing multi-template modelling to encompass the entire protein sequence [21]. Additionally, loop refinement was performed. The resulting structure was validated using the SAVES server, utilizing the ERRAT and PROCHECK utilities for validation [22, 23]. The ProSA program (Protein Structure Analysis), which provides quality scores for a protein based on known structures from X-ray analysis, NMR spectroscopy, and theoretical calculations, was also employed for structural validation [24]. Mutant protein structures were generated using the mutagenesis utility in PyMol. Subsequently, all structures underwent energy minimization using SPDBV 4.10 with the GROMOS96 force-field [25].

2.4 Protein-protein docking of PCSK9 and LDLR

The binding site of PCSK9 has been localized to the epidermal growth factor-like repeat A (EGF-A) domain of the Low-density lipoprotein receptor (LDLR). The three-dimensional structure of the EGF-A domain of LDLR was obtained from the Protein Data Bank (PDBID: 3BPS) [26]. To investigate the impact of mutations on protein-protein interactions (PPI), both wildtype and mutant PCSK9 structures were docked with the LDLR protein using the HawkDock server [27]. The docking process aimed to comprehend alterations in the protein-protein interactions resulting from mutations. The binding sites of the proteins, derived from the 3BPS structure, were utilized in the docking experiments. HawkDock employed the ATTRACT docking algorithm, HawkRank scoring function, and MM/GBSA free energy decomposition to evaluate the docked complexes. The protein-protein interactions were further analyzed and assessed using the PDBsum server [28].

2.5. Molecular dynamics and simulation analysis

The structural stability and flexibility of both wildtype and mutant protein-protein complexes were evaluated through molecular dynamics and simulation (MDS) analysis. Molecular dynamics simulations were conducted using the GROMACS 2020.6 simulation package, employing the Gromos53a6 force field [29]. The modelled wildtype (WT) and mutant forms of PCSK9 were placed at the centre of a cubic box and solvated with the SPC (single point charge) water model. The solvated systems were neutralized with sodium counter ions and energy minimized using the steepest descent algorithm over 50,000 steps. During this process, Coulombic and van der Waals interactions were maintained with cutoff values of 0.9 nm and 1.0 nm, respectively. Subsequently, the systems were equilibrated under NVT and NPT ensembles using the Leap-frog integrator for 100 ps, while Berendsen temperature coupling and pressure coupling were maintained at 300 K and 1 bar, respectively. Long-range electrostatics were handled using Particle-Mesh Ewald (PME), and short-range electrostatics were managed by Coulombic and van der Waals interactions, both with a cutoff of 1.0 nm and a time step of 10 fs. The LINCS algorithm was employed to preserve molecular geometry. Following proper equilibration, the systems underwent 100 ns of MD simulations at 300 K, sampling conformations at 2 ps intervals.

2.6. Binding Free Energy and Trajectory Analysis

The binding free energies of the protein-protein complexes were calculated using the Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) method. This approach combines molecular mechanics energies with an empirical solvation model to estimate the free energy of binding, providing insights into the stability and interactions of the complexes. The MM/GBSA calculations were performed using the software package available in conjunction with our molecular dynamics simulations, allowing for a comprehensive assessment of binding affinities. This method is particularly effective for evaluating the contributions of different energy components, such as van der Waals, electrostatic, and solvation energies, to the overall binding process.

The resulting trajectories of wildtype and mutant were taken for further analysis using the in-built functions in GROMACS, which include gmx_rmsd, gmx_rmsf, gmx_sas, gmx_rg, g_sham, g_anaeig and g_covar.

2.7. Statistical Analysis

To quantitatively assess the significance of observed differences between the wild-type and mutant PCSK9-LDLR complexes, comprehensive statistical analyses were performed on the data extracted from the molecular dynamics trajectories and binding free energy calculations.

For parameters derived from the MD simulations—including Root Mean Square Deviation (RMSD), Root Mean Square Fluctuation (RMSF), Radius of Gyration (Rg), and Solvent Accessible Surface Area (SASA)—the trajectories from the stable simulation period (50-100 ns) were used. The time-series data for each system were treated as independent observations. The normality of the data distribution for each parameter was verified using the Shapiro-Wilk test. Since the data for each system were collected from a single, long simulation, we employed a sampling-based approach by dividing the stable trajectory into multiple independent blocks to ensure statistically robust comparisons.

An independent samples t-test was used for pairwise comparisons between the wild-type and each mutant complex. For global comparisons across all systems (wild-type and the three mutants), a one-way Analysis of Variance (ANOVA) was conducted, followed by a post-hoc Tukey's test to identify which specific pairs were significantly different. The binding free energies calculated using the MM/GBSA method for the top-ranked docking poses were

compared using a similar t-test approach.

All statistical tests were performed using GraphPad Prism software (version 9.0). Data throughout the manuscript are presented as the mean ± standard deviation (SD). A p-value of less than 0.05 was considered statistically significant, with the following conventions: *p < 0.05, **p < 0.01, ***p < 0.001, and ****p < 0.0001.

3. Results

3.1. Variant Retrieval and Initial Annotation

Genetic variants associated with reduced serum low-density lipoprotein cholesterol (LDL-C) levels have been established as protective against coronary heart disease (CHD). Notably, protection is conferred by rare loss-of-function (LoF) and protein-truncating variants within canonical lipid-metabolism genes, including APOB and PCSK9. Through an extensive literature survey, 89 LoF variants of PCSK9 were identified, with 86 of them documented in dbSNP (Table S6). These variants comprised 31 frameshift, 18 stop_gained, 10 splice donors, 5 splice acceptors, 24 missense variants, and 1 in-frame deletion. The variant classes encompassed 59 substitutions, 21 deletions, 7 duplications, and 2 insertions. Comprehensive analysis revealed a higher predisposition of LoF variants to frameshift mutations leading to amino acid substitutions. Upon examining the variant locations within protein domains, they were found in domains such as non_cytoplasmic_domain, Peptidase S8 propeptide/proteinase inhibitor I9 domain, Proprotein convertase subtilisin/kexin type 9, C-terminal domains 1, 2, and 3, Proteinase K-like catalytic domain, Peptidases_S8_PCSK9_ProteinaseK_like domain, Peptidase S8/S53 domain, and Subtilase family. Notably, the Peptidase S8/S53 domain contained the highest number of variants, approximately 21 in total (S. Table 1).

3.2. Functional analysis of missense variants

The functional impact of 24 missense variants in PCSK9 was assessed through sequence analysis using tools like SIFT, SNAP2, and Align-GVGD (Table 1). SIFT, which predicts adverse effects upon amino acid substitution through sequence homology, identified 12 variants as tolerated and 12 as not tolerated. SNAP2 indicated that 13 variants caused changes in molecular function, while 11 had a neutral effect. Align-GVGD graded variants with discrete C-scores based on biochemical variation among amino acids measured by Grantham Deviation (GD). Approximately 14 variants were classified as C65, 5 as C55, 1 as C35, 3 as C25, and 5 as C15, all indicating biochemical variations in the protein. Among these, the 14 variants with C65 were most likely to interfere with protein function. Comparing the analysis tools, 8 variants were predicted by both SIFT and Align-GVGD to impact the protein's primary function. Notably, variants rs146471967, rs757143429, rs148195424, and rs1553137543 were predicted to alter the protein by all tools. The consistent findings across these tools suggest that evolutionarily conserved regions are less tolerant of mutations, indicating that mutations in these regions are more likely to affect protein function.

3.3. Thermal stability analysis

Certain missense mutations can influence the thermal stability, unfolding, and aggregation of proteins under physiological conditions. Therefore, we examined the thermal stability of 24 missense variants using I-mutant 3.0, MUPro, DDGun, and DUET (Table 2). These tools calculated the change in Gibbs free energy resulting from the difference between the unfolding Gibbs free energy of the mutated protein and the unfolding free energy of the native protein (in kcal/mol). According to the predictions, I-mutant 3.0 identified 12 neutral variants, 7 variants with a significant decrease in free energy, and 5 variants with a notable increase in free energy. DDGun, which also employs ΔΔG values to assess the functional impact of variants, predicted 8 neutral variants, 4 variants with increased free energy, and 12 variants with decreased free

Table 1. Functional impact as predicted.

Variant	Substitutions	SIFT	SNAP2	Align-GVGD
rs11591147	R46L	Tolerated	Effect	Class C65
rs151193009	R93C	Tolerated	Effect	Class C65
rs369067856	R104C	Not Tolerated	Neutral	Class C65
	G106R	Not Tolerated	Neutral	Class C65
rs775988212	V114A	Tolerated	Effect	Class C55
rs143117125	N157K	Tolerated	Neutral	Class C65
rs778617372	Q219E	Tolerated	Effect	Class C25
	A239D*	Tolerated	Neutral	Class C65
rs149489325	G236S	Tolerated	Effect	Class C55
rs146471967	H391N	Not Tolerated	Effect	Class C65
rs28362263	A443T	Tolerated	Effect	Class C55
rs746115963	S462P	Not Tolerated	Neutral	Class C25
	Q554E*	Not Tolerated	Neutral	Class C25
rs755750316	P616L	Not Tolerated	Neutral	Class C65
rs762298323	S668R	Not Tolerated	Neutral	Class C65
rs145886902	E57K	Tolerated	Neutral	Class C55
rs757143429	R434W	Not Tolerated	Effect	Class C65
rs533273863	P174S	Tolerated	Neutral	Class C65
rs747072726	R105Q	Not Tolerated	Neutral	Class C35
rs72646508	L253F	Tolerated	Effect	Class C15
rs766999045	R96L	Tolerated	Effect	Class C65
rs148195424	R237W	Not Tolerated	Effect	Class C65
rs1553137543	N354I	Not Tolerated	Effect	Class C65
rs777300852	A522T	Not Tolerated	Effect	Class C55

Table 2. Thermal stability analysis of the variants.

Variant	Substitutions	SIFT	SNAP2	Align-GVGD
rs11591147	R46L	Tolerated	Effect	Class C65
rs151193009	R93C	Tolerated	Effect	Class C65
rs369067856	R104C	Not Tolerated	Neutral	Class C65
	G106R	Not Tolerated	Neutral	Class C65
rs775988212	V114A	Tolerated	Effect	Class C55
rs143117125	N157K	Tolerated	Neutral	Class C65
rs778617372	Q219E	Tolerated	Effect	Class C25
	A239D*	Tolerated	Neutral	Class C65
rs149489325	G236S	Tolerated	Effect	Class C55
rs146471967	H391N	Not Tolerated	Effect	Class C65
rs28362263	A443T	Tolerated	Effect	Class C55
rs746115963	S462P	Not Tolerated	Neutral	Class C25
	Q554E*	Not Tolerated	Neutral	Class C25
rs755750316	P616L	Not Tolerated	Neutral	Class C65
rs762298323	S668R	Not Tolerated	Neutral	Class C65
rs145886902	E57K	Tolerated	Neutral	Class C55
rs757143429	R434W	Not Tolerated	Effect	Class C65
rs533273863	P174S	Tolerated	Neutral	Class C65
rs747072726	R105Q	Not Tolerated	Neutral	Class C35
rs72646508	L253F	Tolerated	Effect	Class C15
rs766999045	R96L	Tolerated	Effect	Class C65
rs148195424	R237W	Not Tolerated	Effect	Class C65
rs1553137543	N354I	Not Tolerated	Effect	Class C65
rs777300852	A522T	Not Tolerated	Effect	Class C55

*Not reported in dbSNP; High impact variants are highlighted based on the tool results.

energy. DUET, utilizing predictions from two complementary methods, mCSM and SDM, pinpointed 18 destabilizing variants ($\Delta\Delta G$<-2 kcal/mol) and 6 stabilizing variants. The comparative analysis of the predictions from these *in-silico* tools indicated rs369067856, rs146471967, rs757143429, and rs1553137543 to have a significant impact on the molecular function by altering the thermal stability of the protein.

3.3. Wildtype and mutant solvent interaction

Comprehending the exposure of a residue within a folded protein's structure is vital for delineating the protein folding core and pinpointing residues that interact with solvents and other molecules in both natural and artificial settings. DeepREx-WS, besides assessing solvent exposure, computed several parameters like conservation, hydrophobicity, and flexibility (Table 3). Out of the 24 variants scrutinized, 9 variants displayed no change in conservation (Conservative Index (CI = 0), while 15 variants were moderately conserved (0 < CI < 1). Although there were slight differences in CI among these 15 variants, the CI remained consistent in the 9 variants when compared to the wildtype. The hydrophobicity profile of wildtype residues indicated 17 hydrophilic and 7 hydrophobic residues. Five variants (rs11591147, rs369067856, rs757143429, rs766999045, and rs1553137543) transitioned from hydrophilic to hydrophobic nature. The G106R variant, not documented in dbSNP, exhibited increased hydrophilicity, whereas rs533273863 displayed decreased hydrophilicity. Furthermore, rs755750316 demonstrated heightened hydrophobicity, while rs777300852 showed diminished hydrophobicity. In three variants (rs11591147, rs746115963, and rs757143429), solvent-exposed residues became buried, whereas buried residues of rs778617372, A239D (not reported in dbSNP), and rs28362263 became exposed. A comparison of flexibility between wildtype and mutant residues identified 5 variants with reduced flexibility and 2 variants with increased flexibility. The comprehensive analysis revealed that the variants modified their solvent accessibility and flexibility, thereby impacting the protein's structure and function.

3.4. Protein-Protein interaction

Based on the outcomes obtained from prediction tools and existing literature, three specific PCSK9 variants were selected for investigation concerning their impact on binding with the EGF-A domain of LDLR. These variants included two single mutants, N354I (rs1553137543) and H391N (rs146471967), as well as one double mutant, R104C-V114A (rs369067856 and rs775988212).

The structural validation of the modelled wildtype protein is detailed in the table (Table 4). Mutant protein structures were predicted by altering the corresponding amino acid residues in the wildtype PCSK9 protein structure (Fig. 1). Protein-protein docking experiments between the wildtype, single mutants, and double mutants of PCSK9 with LDLR protein were conducted using the HawkDock server. Ten models were generated for the wildtype and mutant PCSK9 protein complexes with LDLR, which were then re-ranked using MM-GBSA based on binding free energies. The best-scoring model was analysed to gain insight into the contributing energies and key residues. Additionally, per-residue free energy contributions were studied.

Comparative analysis of the MM-GBSA contributing energies revealed that van der Waals energy was the most significant, followed by electrostatic and solvation energy. Overall energy comparison indicated that the double mutant (R104C-V114A) exhibited higher values in all energy terms compared to wildtype and mutant protein complexes. Two key residues, PHE210 and GLN250, were found to be crucial in all mutant complexes. Some key residues identified in the mutants were shared between one of the other mutant protein complexes (Tables 5 & 6).

Table 3. Solvent accessibility, flexibility and conservation of the LOF variants.

Variant	Substitution	Wildtype				Mutant			
		Hydrophobicity	Conservation	Flexibility	Exposure	Hydrophobicity	Conservation	Flexibility	Exposure
rs11591147	R46L	-0.64	0.0	3	Exposed	1.02	0.0	3	Buried
rs151193009	R93C	-1.54	0.18857	0	Exposed	-0.14	0.19121	1	Exposed
rs369067856	R104C	-1.16	0.19223	2	Exposed	0.24	0.19438	2	Exposed
	G106R*	-1.38	0.20901	1	Exposed	-2.2	0.19406	1	Exposed
rs775988212	V114A	0.88	0.18006	0	Buried	0.4	0.1748	0	Buried
rs143117125	N157K	-1.14	0.26946	0	Exposed	-1.22	0.2723	1	Exposed
rs778617372	Q219E	-2.04	0.19199	1	Buried	-2.04	0.19165	2	Exposed
	A239D*	-0.48	0.25979	1	Buried	-1.54	0.2527	2	Exposed
rs149489325	G236S	-1.0	0.33834	0	Buried	-1.08	0.31602	0	Buried
rs146471967	H391N	1.28	0.39544	0	Buried	1.22	0.3834	0	Buried
rs28362263	A443T	2.0	0.0	0	Buried	1.5	0	0	Exposed
rs746115963	S462P	0.22	0.0	0	Exposed	0.06	0	1	Buried
	Q554E*	-1.62	0.0	4	Exposed	-1.62	0	4	Exposed
rs755750316	P616L	0.54	0.0	4	Exposed	1.62	0	4	Exposed
rs762298323	S668R	-1.24	0.0	4	Exposed	-1.98	0	3	Exposed
rs145886902	E57K	-1.38	0.0	4	Exposed	-1.46	0	4	Exposed
rs757143429	R434W	-0.7	0.0	1	Exposed	0.02	0.0	1	Buried
rs533273863	P174S	-2.12	0.16286	3	Exposed	-1.96	0.15643	4	Exposed
rs747072726	R105Q	-1.78	0.19878	2	Exposed	-1.58	0.20095	2	Exposed
rs72646508	L253F	0.5	0.29023	0	Buried	0.3	0.3002	0	Buried
rs766999045	R96L	-0.82	0.18351	0	Exposed	0.84	0.18158	0	Exposed
rs148195424	R237W	-1.48	0.23411	1	Exposed	-0.76	0.22975	1	Exposed
rs1553137543	N354I	-0.44	0.27113	0	Buried	1.16	0.23159	0	Buried
rs777300852	A522T	2.2	0.0	0	Buried	1.7	0	0	Buried

*Not reported in dbSNP

Table 4. Template and structure validation statistics of the modelled proteins.

Template ID	Query Coverage (%)	Identity (%)	Procheck (%)				ERRAT*	ProSA**
			A	B	C	D		
2P4E	100	99.71						
6U2F	100	99.71	91	7.5	0.9	0.6	70.125	-10.13
3BPS	78	99.63						

[A]Most favoured regions, [B]Additionally allowed regions, [C]Generously allowed regions, [D]Disallowed regions, *Overall quality, **Z score

Fig. 1. Cartoon representation of PCSK9 protein. The single mutants N354I, H391N and double mutant R104C-V114A are shown in ball and stick.

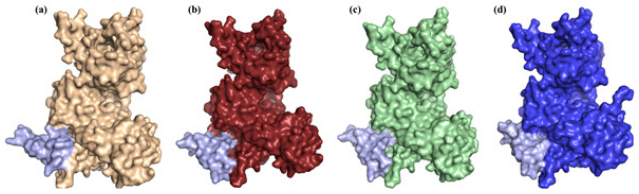

Fig. 2. Protein-protein docked complex of PCSK9 and LDLR protein. (a) wild type (b) N354I (c) H391N and (d) double mutant (R104C - V114A). The protein-protein docked complex is represented in surface representation. The PCSK9 protein is coloured in wheat, brick red, green and dark blue for the wildtype, N354I, H391N and double mutant, respectively. The light blue surface represents LDLR protein.

Upon analysing the protein-protein complexes (Fig. 2), it was observed that the number of hydrogen bonding interactions decreased in the mutants compared to the wildtype. However, an increase in non-bonding interactions was observed in all mutants. Salt bridges were only present in the mutant structures. Examination of the conservation of interacting residues revealed that the single mutants shared the same non-bonding interacting residues. Hydrogen bonding analysis showed that SER165 of PCSK9 wildtype formed a hydrogen bond with ASN9 of LDLR, whereas SER165 of PCSK9 single mutants formed a hydrogen bond with ARG38 of LDLR. Similarly, ASP368 of PCSK9 wildtype was involved in hydrogen bonding interactions with ASN9 and ASN10 of LDLR, while ASP368 of PCSK9 single mutants participated in hydrogen bonding with ARG38 of LDLR. SER312 and SER313 of wildtype PCSK9 formed non-bonded interactions with ASN9 of LDLR, while SER312 and SER313 of all mutants were engaged in non-bonded interactions with ARG38 of LDLR. Although the double mutant preferred the same binding cavity, it did not exhibit conservation in the interacting residues involved in hydrogen bonding. Both single mutants formed salt bridges between ASP368 of PCSK9 and ARG38 of LDLR, while the double mutant formed a salt bridge with ASP367 of PCSK9 and ARG39 of LDLR (S. Tables 2-5).

3.5. Molecular dynamics and simulation analysis

Molecular dynamics simulations are considered a reliable method for probing the conformational stability of wildtype and mutant complexes. In this regard, an appropriate complex with the lowest binding energy of wild type and mutant was chosen as the input files for MD computation to analyse the conformational stability and comprehend the structural aspects of the complex. MDS involves

Table 5. MM-GBSA energy contribution.

Variant	Score	VDW*	ELE**	GB[&]	SA[&&]	Binding Energy (kcal/mol)
Wildtype	-43.68	-82.23	-23.23	94.24	-10.59	-21.81
N354I	-42.05	-76.22	-72.50	136.13	-10.40	-22.98
H391N	-48.39	-76.37	-74.41	142.59	-10.56	-18.74
R104C - V114A	-26.88	-97.96	-232.35	307.43	-12.23	-35.12

*van der Waals interaction energy, **Electrostatic energy, [&]Polar solvation energy, [&&]Solvation energy.

Table 6. Key residues identified by MM-GBSA method.

Rank	Wildtype		N354I		H391N		R104C - V114A	
	Chain A*	Chain B[+]	Chain A*	Chain B[+]	Chain A*	Chain B[+]	Chain A*	Chain B[+]
1	GLN-322	ARG-38	ARG-235	PHE-32	PHE-4	THR-3	PHE-162	PRO-29
2	VAL-320	GLY-12	LEU-288	PRO-29	HIE-166	PRO-29	LEU-294	VAL-16
3	SER-165	ASN-9	PHE-156	SER-14	PHE-156	LEU-7	GLN-328	LEU-7
4	ASP-314	ARG-39	GLN-196	ASN-9	TYR-233	PHE-32	SER-329	ASN-9
5	PHE-156	SER-14	TYR-233	ASN-10	SER-165	LEU-20	ARG-12	SER-14
6	SER-321	THR-3	ARG-297	VAL-16	LEU-264	GLY-2	VAL-326	ASN-10
7	SER-323	TYR-24	TRP-12	GLN-33	AL8	ASN-4	GLY-295	ARG-39
8	ILE-309	AL36	LEU-264	LEU-7	HIE-5	ILE-22	SER-327	PHE-32
9	HIE-166	GLN-37	HIE-5	ARG-39	GLN-196	VAL-16	HIE-11	ASN-18
10	GLN-196	CYS-13	MET-410	HIE-15	PHE-90	GLN-37	ILE-315	LEU-20

*PCSK9, [+]LDLR

performing intense force field calculations for each atom in a molecular system, followed by incorporation phases that advance the atoms' locations and nature. Several calculations were used to examine wildtype and mutant structural alterations through RMSD, RMSF, Rg, SASA, PCA and FEL computations. The xmgrace module was used to generate graphs from the trajectories of all the systems.

3.6. Root Mean Square Deviation

The deviations in a molecule's backbone, from its initial structural conformation to its final position, are estimated during simulation with the GROMACS command-line interface gmx_rms. The deviations of the interactions were evaluated using 100 ns simulated trajectories. The divergence observed during simulation measures the structural stability of the wild type (WT) and mutants (protein-protein complexes). A plot displaying RMSD (nm) vs. time (ns) for the WT and mutant protein-protein complexes is shown in Fig. 3a. Here, the average RMSD of WT, single mutants N354I and H391N and double mutants are 0.33 nm (WT), 0.56 nm, 0.53 nm and 0.52 nm, respectively. The RMSD plot showed that the deviations in the wildtype protein complex were lower compared to the mutant complexes. The N354I mutation exhibited the highest deviation, followed by the H391N mutation and the double mutant. Substitution of amino acid residues in the mutant systems can induce perturbations and conformational rearrangements, leading to elevated RMSD values in the mutant complexes. Since the average RMSD value of WT is less than 2Å, it can be deemed to be stable, validating the structure's equilibration protocol. To further characterize and validate the structure, in-depth investigations were carried out as detailed in the subsequent section.

3.7. Root Mean Square Fluctuation

To gain more insight into the stability and dynamics of residues in the wildtype and mutant complexes, residue-based root mean square fluctuation (RMSF) was assessed. The RMSF function was used to evaluate the conformational changes of the structural residues. Analysis of RMSF plots revealed flexible regions of the residues involved in the WT and mutants over the time scale of MD trajectories. Generally, a more considerable RMSF value indicated loosely organized loops (higher fluctuations) or terminal ends, while a lower value suggested well-structured areas. Fig. 3b. displays the RMSF plot of wild type and mutants. The RMSF plot revealed higher fluctuations in the mutant protein complexes compared to the wildtype. Specifically, N354I exhibited higher residual fluctuations than the wildtype (Table 7). On closer observation, the higher residual fluctuations were exhibited by the residues located in the loop regions.

3.8. Radius of Gyration

The Rg is a criterion used to quantify the structural rigidity of complex biological systems concerning their geometry and topology. To conclude whether the wild type and mutants were steadily folded or unfolded during MD simulation and to arrange structural convergence, the radius of gyration was computed against time. Fig. 4a shows the Rg curves as a function of time in WT and mutant complexes. The Rg plot showed stable values for the WT until 90 ns, after which a decrease was observed. In the case of N354I mutation, the Rg plot showed an increased value after 72ns, reaching approximately 2.81nm by the end of the simulation. The Rg value of H391N and double mutant fluctuated primarily and increased after 36ns. Finally, a minimal Rg value of wildtype shows that the complex remained compact throughout the simulation time. The increased Rg value observed in the mutant complexes suggests that the substitutions of residues have reduced the compactness of the PCSK9 protein, likely due to alterations in the protein's structure and folding.

3.9. Solvent Accessible Surface Area

Solvent Accessible Surface Area (SASA) analysis was utilized to measure the exposure of amino acid residues to the surrounding environment. An elevated SASA value indicates a more dispersed protein structure, while a decreased SASA value suggests a more compact structure. The SASA values of the WT and mutant protein complexes were examined to anticipate the influence of amino acid substitutions on protein structure. A plot showing the SASA value is shown in Fig. 4b. The increased SASA values in the N354I and double mutant complexes suggest a greater solvent-accessible area, indicating a larger exposed surface. Conversely, the lower SASA value for H391N implies reduced accessibility to the solvent during the 100 ns simulation. However, towards the end of the

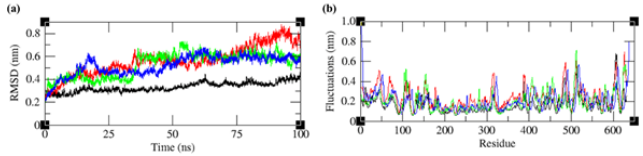

Fig. 3. Plots illustrating (a) the RMSD (b) RMSF of PCSK9-LDLR docked wild type and mutant complex. Black – Wild type, Red – N354I, green – H391N and blue double mutant (R104C - V114A).

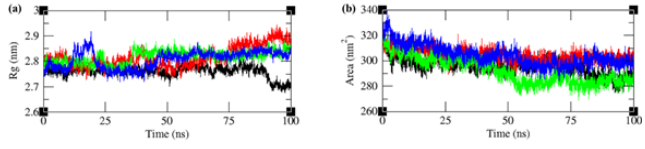

Fig. 4. Plots illustrating (a) the Rg (b) SASA of PCSK9-LDLR docked wild type and mutant complex. Black – Wild type, Red – N354I, green – H391N and blue double mutant (R104C - V114A).

Table 7. Average values were computed for Rg, RMSF, and RMSD of wildtype and mutants.

Sl. No.	Average Value	Wild Type	N354I	H391N	R104C - V114A
1	RMSD*	0.33	0.56	0.53	0.52
2	RMSF*	0.17	0.24	0.20	0.21
3	Rg*	2.07	2.81	2.81	2.8
4	SASA**	295.73	304.44	292.39	304.21

Units *nm, +nm**

simulation period, the WT and H391N mutant complex exhibited comparable behaviour.

3.10. Principal Component Analysis

Although the degree of freedom of the receptor is considerable, most of its motion is concentrated on a minimal number of degrees of freedom, in which the motion is closely related to the protein's function. Principal component analysis (PCA) can separate the subspace of receptor, making it possible to understand the functional movement patterns of proteins more clearly. Based on the MD trajectories of all the complexes, the covariance matrix is constructed for Cα atom, and the covariance matrix is then diagonalized into eigenvalues and corresponding eigenvectors. The simulated trajectory is projected onto the basis vectors, and the two projection coordinates with the most significant deviation are stacked to visually see the amplitude of functional motion. The comparison of the first two principal components of the WT with the mutant protein complexes indicated that conformational states of the mutant protein complexes were significantly changed. As shown in Fig. 5, the PC2 versus PC1 plots for all protein complex systems clearly indicate that the mutant protein complex conformers visit a large conformational space when compared to WT. This indicates that mutations in these proteins have caused conformational changes. However, the presence of overlap in the conformational space indicates that the protein does revisit the same state during the simulation, even though it undergoes conformational changes.

The mutations lead to a general enhancement of collective motions in specific proteins, effectively amplifying the Cα movements within the protein structure. Consequently, these mutations augment the overall flexibility of the proteins. This implies that the mutations induce alterations in the dynamics of the mutated proteins, potentially contributing to their proper functionality. Moreover, the fluctuations observed in each protein can be considered essential for maintaining stability throughout molecular dynamics simulations over time. The Principal Component Analysis (PCA) results align with findings from RMSD, RMSF, Rg, and SASA analyses.

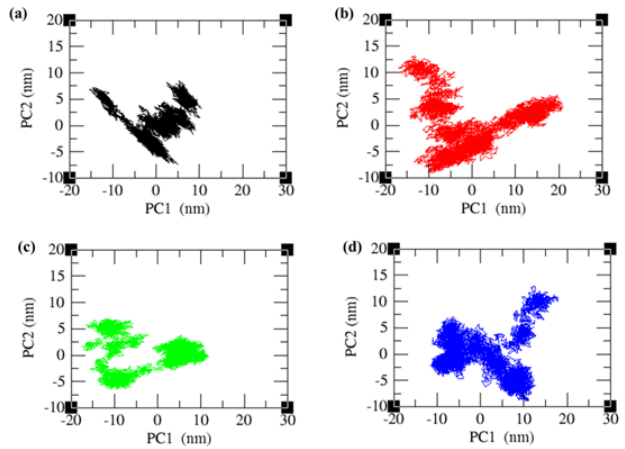

Fig. 5. Principal component analysis of the wild type and mutant protein complexes. The plots show the projection of PC2 vs PC1 of (a) wild type (b) N354I (c) H391N and (d) double mutant (R104C - V114A), respectively.

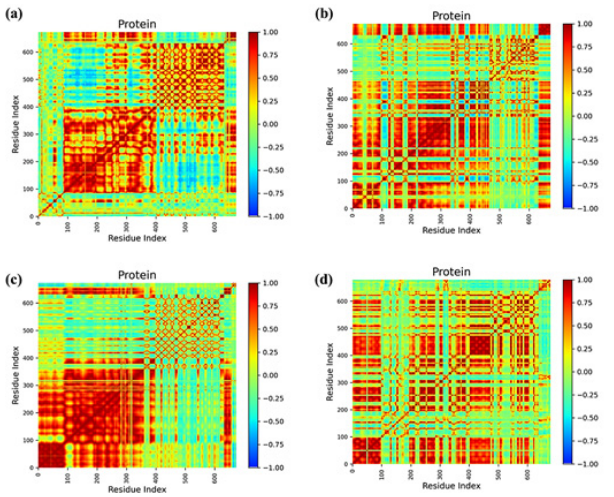

Fig. 6. Comparison of cross-correlation matrices of wild type and mutant protein complexes. Cross correlation matrix of Cα atoms during 100 ns simulation for (a) wild type (b) N354I (c) H391N and (d) double mutant (R104C - V114A), respectively. The range of motion is denoted by different colours in the panel. Red colour indicates positive correlation and blue colour indicates anti-correlation motions.

3.11. Dynamical cross-correlated map (DCCM) analysis

To further explore the impact of mutations on conformational dynamics, Differential Cross-Correlation Mapping (DCCM) analysis was employed to assess Cα atom fluctuations during the 100-ns simulation. This analysis allowed for the examination of correlated motions among different regions of each protein complex. The cross-correlation values were normalized and ranged from -1.0 to +1.0 (from dark blue to green to red) as depicted in Fig. 6. Positive correlations (red regions) indicated residues moving in the same direction, while negative correlations (blue regions) signified residues moving in opposite directions. The intensity of the colour reflected the strength of the correlation, with deeper shades indicating stronger positive or negative correlation. Regions in green, ranging from -0.25 to 0.25, were considered to have low correlation. Diagonal points represented Cα atoms of the same residue along both axes, with diagonal elements indicating maximum correlation.

The cross-correlation matrix of Cα displacement revealed intricate patterns of correlated and anti-correlated motions in both the wild-type and all mutant protein complexes. Notably, in all mutant protein complexes, there was an increase in correlated motions and a decrease in anti-correlated motions compared to the WT. The mutated proteins exhibited deeper shades of red, indicating regions with higher correlation, suggesting predominantly global motions in these protein complexes. In most mutant protein complexes, nearly all residues demonstrated concerted motions, moving in a correlated fashion. The principal components and dynamic cross-correlation maps further highlighted that the mutations alter the major motions of the proteins, potentially enhancing their dynamic behaviour. These findings underscore the substantial impact of the mutations on the conformational flexibility of proteins.

3.12. Free-energy calculations

Conformational transitions and structural changes in protein-protein complexes were investigated through Free

Energy Landscape (FEL) analysis using a comprehensive conformational sampling procedure. The top two principal components, PC1 and PC2, were identified as significant contributors and were utilized to construct FEL contour plots, enabling the monitoring and comprehension of the native and metastable states within each system.

In general, Gibbs free energy-based contour plots with single and multiple conformation clusters indicate smaller and larger conformational changes, respectively. Fig. 7 presents the FEL values for the WT and mutant protein complexes, respectively, with deeper colours indicating lower energy levels. For both the WT and double mutant protein complexes, only two main free energy wells/basins were observed in the global free energy minimum region. Conversely, single mutants N354I and H391N protein complexes exhibited three and one main free energy wells in the global free energy minimum region, respectively.

A comprehensive comparison of the FEL values between WT and mutant protein complexes revealed that mutant systems spanned larger ranges of PC1 and PC2, displaying a more rugged free energy surface than the WT protein complex. Notably, the FEL plots of N354I and double mutants showcased a greater number of local free-energy minima, while H391N exhibited a decrease in local free-energy minima, indicating conformational rearrangements induced by the mutations.

In this study, we identified 89 loss-of-function (LoF) variants in the PCSK9 gene that are associated with reduced serum LDL cholesterol (LDL-C) levels, providing significant insights into their protective role against coronary heart disease (CHD). Our findings align with existing literature that supports the hypothesis that such variants in lipid-metabolism genes confer cardiovascular protection. For instance, prior studies have documented similar LoF mutations in PCSK9 and their correlation with decreased LDL-C levels and reduced risk of CHD, reinforcing the clinical relevance of our work.

Our comprehensive analysis of variant classes and their locations within critical protein domains, particularly the Peptidase S8/S53 domain, offers a more nuanced understanding of how these mutations disrupt PCSK9's function. Experimental data, such as structural studies and biochemical assays, have previously highlighted the importance of these domains in mediating interactions with LDL receptors. By situating our findings within this context, we enhance the current knowledge base and provide a foundation for future research aimed at elucidating the precise mechanisms by which these variants influence lipid metabolism. Moreover, our functional analyses, employing tools like SIFT and DDGun, identified specific missense variants that significantly impact thermal stability and solvent accessibility. These computational predictions are supported by experimental observations in other studies, suggesting that alterations in thermal stability can lead to functional consequences for protein interactions. Our molecular dynamics simulations further demonstrate that these variants increase the flexibility of the PCSK9 protein, corroborating findings from similar studies that have linked enhanced protein dynamics to altered functional outcomes. Overall, our work not only confirms previous associations between LoF variants in PCSK9 and cardiovascular health but also provides novel insights into the structural and dynamic consequences of these mutations. This integrated approach, combining computational analyses with established experimental data, informs ongoing discussions about the genetic factors influencing lipid metabolism and their implications for therapeutic interventions. Future research, particularly functional validation of these variants, will be crucial in translating these findings into clinical applications and understanding the broader implications for genetic counselling and patient management.

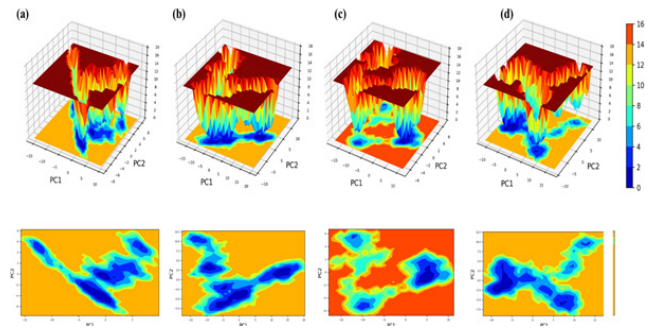

Fig. 7. 3D and 2D Free energy contour map constructed based on PC1 and PC2 for (a) wild type (b) N354I (c) H391N and (d) double mutant (R104C - V114A), respectively. Deeper colour area in the maps indicates lower energy. The colour bar represents the relative free-energy value in kcal mol^{-1}.

This study successfully addresses the objective of investigating how loss-of-function (LOF) variants in PCSK9 contribute to the reduction of LDL cholesterol (LDL-C) levels and their potential relationship with cardiovascular disease (CVD). By identifying 89 LOF variants and conducting extensive functional analyses, we established that these variants are predominantly located in critical protein domains, including the Peptidase S8/S53 domain, which is essential for PCSK9's interaction with LDL receptors. Our findings indicate that many of these LOF variants significantly alter thermal stability, solvent accessibility, and protein-protein interactions, which collectively enhance the likelihood of lower LDL-C levels in carriers. Furthermore, molecular dynamics simulations revealed that these variants lead to increased flexibility and alterations in structural dynamics, providing a mechanistic understanding of how they might confer protection against CVD. Overall, the results highlight the pivotal role of these genetic variants in modulating cholesterol metabolism, thereby linking them to reduced LDL-C levels and a potentially lower risk of coronary heart disease.

The possible future research avenues that may stem from our study on genetic variants in PCSK9 and their relationship with LDL-C levels and coronary heart disease. Here are the key proposed directions: (i) Functional Validation of Variants, (ii) Exploration of Additional Genetic Variants, (iii) Functional Impact on Protein-Protein Interactions, (iv) Pathway Analysis and Network Biology, (v) Therapeutic Implications, and (vi) Clinical Implications and Genetic Counselling, etc. We believe that these outlined directions will provide significant insights and practical applications stemming from our findings.

In light of the limitations we encountered regarding ACMG-based analyses, it is essential to highlight the rationale behind our current approach to assessing the clinical effects of loss-of-function (LOF) variants. While we acknowledge that comprehensive ACMG evaluations would provide a deeper understanding of the clinical impli-

cations, our focus has instead been on leveraging available computational analyses and correlating these findings with clinically reported phenotypes. This strategy allows us to draw preliminary connections between the identified variants and their potential impact on patient health, even in the absence of extensive clinical data. By discussing these computational results and their relevance to existing literature on phenotypic outcomes, we aim to provide a meaningful context for our findings and lay the groundwork for future investigations that may include more detailed clinical assessments. This approach not only addresses the current feasibility constraints but also underscores the importance of integrating computational insights with clinical observations in the ongoing exploration of genetic variant effects.

4. Discussion

Our computational analysis provides a mechanistic framework for understanding how loss-of-function (LOF) variants in PCSK9 can translate into cardiovascular benefits. The hepatic secretory protein PCSK9 binds to the low-density lipoprotein receptor (LDLR) and escorts it to lysosomal degradation. By lowering LDLR abundance on hepatocytes, PCSK9 reduces LDL clearance and raises circulating LDL-cholesterol (LDL-C). The genetic discovery that PCSK9 mutations cause autosomal-dominant hypercholesterolemia established the centrality of the pathway to lipoprotein metabolism. Subsequent biochemical studies and reviews have consolidated PCSK9's role as a gatekeeper of LDLR turnover [30, 31]. Our docking and molecular dynamics (MD) results support this. We observed reduced binding energies, altered contact residues, preservation of the cavity but remodeling of interactions, and increased flexibility/solvent exposure. These findings are consistent with known features of the PCSK9–LDLR interface. PCSK9 interacts with the EGF-A module within the LDLR EGF homology domain. Crystal and biochemical data show that specific polar and hydrophobic contacts at this surface form a productive complex. In this structural context, LOF substitutions that destabilize local packing, disrupt hydrogen-bond networks, or increase loop mobility are expected to weaken affinity. This can occur without relocating the binding site precisely, as observed in our in-silico work [32-34].

Functionally, reduced affinity of LOF variants provides a plausible molecular explanation for the human phenotype of lower LDL-C and reduced coronary events. Large population studies have shown that naturally occurring PCSK9 LOF variants are associated with substantially lower lifelong LDL-C and marked reductions in coronary heart disease risk. This highlights a dose–response relationship between diminished PCSK9 activity, enhanced LDLR recycling, and atheroprotection. Our stability and flexibility shifts offer a structural basis for those epidemiologic observations [32,35,36]. These insights dovetail with therapeutic experience. Monoclonal antibodies that block circulating PCSK9 phenocopy LOF genotypes and drive large LDL-C reductions. In the FOURIER outcomes trial, evolocumab reduced cardiovascular events when added to statins. Likewise, siRNA-mediated hepatic PCSK9 silencing with inclisiran achieves sustained LDL-C lowering in Phase 3 clinical trials. This supports the principle that attenuating PCSK9-LDLR engagement is clinically meaningful. Our findings reinforce that weakening the interaction, whether through genetics, antibodies, or gene silencing, converges on the same biology: increased LDLR at the cell surface and enhanced LDL clearance [37,38].

The work also suggests testable hypotheses. First, variants that increase solvent exposure and backbone mobility at the PCSK9-EGF-A interface should display faster dissociation (higher koff) and/or slower association (lower kon) under physiologic and endosomal pH. Surface plasmon resonance or biolayer interferometry, spanning a pH range of 7.4 to 5.5, could validate this [39,40]. Second, the LDLR pathway is sensitive to the endosomal milieu. Experiments in intact hepatocyte systems quantifying LDLR levels, LDL uptake, and receptor degradation kinetics would reveal whether the modeled energetic penalties lead to cellular rescue of LDLR. Third, structural description (e.g., cryo-EM or co-crystallography with EGF-A peptides) of select LOF variants could explain whether altered side-chain rotamers versus larger domain motions drive the affinity loss. Finally, integrating clinical variant data, such as allele frequency and lipid phenotypes, with our per-variant $\Delta\Delta G$ and flexibility metrics could help prioritize variants of uncertain significance [30,32,41]. Limitations warrant consideration. Docking scores and classical MD provide relative, not absolute, affinities and can underrepresent contributions from long-range electrostatics, glycosylation, and conformational selection. Our simulations did not include the full LDLR ectodomain in a membrane/endo-lysosomal context. We also did not explicitly model pH-dependent protonation states known to modulate PCSK9-LDLR binding. Even so, the convergence of independent signals lowers computed binding energies, redistributes contact maps, and increases protein flexibility/solvent exposure. This supports the robustness of the central conclusion: PCSK9 LOF variants induce structural perturbations that weaken LDLR engagement [39,40,42,43].

In summary, the computational picture aligns strongly with human genetics and clinical trials. PCSK9 LOF variants disrupt the fine-tuned interaction with LDLR through effects on stability and dynamics. This leads to decreased affinity, preserved binding topology, and functional protection by enhancing LDLR recycling and LDL-C clearance. These results underscore why PCSK9 remains a validated and tractable target for reducing CVD risk, and provide structural hypotheses to guide the interpretation of variants and the design of next-generation inhibitors.

5. Conclusion

PCSK9 plays a vital role in regulating cholesterol levels in the bloodstream by facilitating the degradation of LDL receptors (LDLR) in the liver. However, the presence of Loss-Of-Function (LOF) variants in the PCSK9 gene results in reduced PCSK9 activity. Individuals carrying PCSK9 LOF variants typically have lower LDL cholesterol levels, reducing their risk of cardiovascular diseases. The analysis of PCSK9 LOF variants using various in-silico tools has provided valuable insights into cholesterol metabolism. The study revealed that LOF variants tend to alter their characteristics, solvent accessibility, flexibility, and free energy, which are crucial for binding with LDLR. Additionally, protein-protein docking and simulation analysis indicated that the mutants became less compact and more exposed to solvent interactions. In comparison to the wildtype PCSK9, only a few key residues contributing

to binding and interaction with LDLR were conserved in the mutants. This study highlights the notable alterations in LDLR degradation caused by LOF variants of PCSK9. These changes suggest that LOF variants could lead to a more stable interaction between LDLR and other regulatory pathways. Consequently, understanding these dynamics may provide new avenues for therapeutic interventions aimed at managing cholesterol levels. Overall, the findings underscore the importance of genetic factors in cardiovascular health and their potential implications for personalized medicine.

Abbreviation

CVD: Cardiovascular disease; **LDLR:** Low-density lipoprotein receptor; **LOF:** Loss-of-function; **GOF:** Gain of function; **SIFT**: Sorting Intolerant from Tolerant; **SVM**: Support vector machine; **MDS:** Molecular dynamics and simulation; **WT**: Modelled wildtype; **PPI**: Protein-protein interactions; **PME**: Particle-Mesh Ewald; **CHD:** Coronary heart disease; **SASA**: Solvent Accessible Surface Area; **PCA**: Principal component analysis; **DCCM:** Differential Cross-Correlation Mapping; **FEL**: Free Energy Landscape.

Conflict of interest

The authors declare that there is no conflict of interest among them.

Consent for publications

The author read & approved the final manuscript for publication.

Ethics approval and consent to participate

This article does not contain any studies with human participants or animals.

Acknowledgement

This study is supported by Prince Sattam bin Abdulaziz University, project number PSAU/2024/R/1445.

Authors' contributions

Tarique M designed the work, analyzed the data, and wrote the final draft. Patel AA and Mallick AK wrote the first draft analysis. Hussien RA, Mirdad TM, Zia-Ul-Sabah, and Rasha Mirdad performed the work and wrote the methodology. Alghawanmeh, IM, and Awadelkarim, MBA, wrote the results and discussion. Kamal MA and Alothaid H helped with the introduction, results, and final manuscript. All authors read and approved the final version.

References

1. Francula-Zaninovic S, Nola IA (2018) Management of measurable variable cardiovascular disease risk factors. Curr Cardiol Rev 14:153-163. doi:10.2174/1573403X14666180222102312
2. Sundararaman SS, Döring Y, van der Vorst EPC (2021) PCSK9: a multi-faceted protein that is involved in cardiovascular biology. Biomedicines 9:793. doi:10.3390/biomedicines9070793
3. Xia XD, Peng ZS, Gu HM, Wang M, Wang GQ, Zhang DW (2021) Regulation of PCSK9 expression and function: mechanisms and therapeutic implications. Front Cardiovasc Med 8:764038. doi:10.3389/fcvm.2021.764038
4. Peterson AS, Fong LG, Young SG (2008) PCSK9 function and physiology. J Lipid Res 49:1152-1156. doi:10.1194/jlr.E800008-JLR200
5. Qiu C, Zeng P, Li X, Zhang Z, Pan B, Peng ZYF, et al. (2017) What is the impact of PCSK9 rs505151 and rs11591147 polymorphisms on serum lipids level and cardiovascular risk: a meta-analysis. Lipids Health Dis 16:111. doi:10.1186/s12944-017-0506-6
6. Benn M, Nordestgaard BG, Grande P, Schnohr P, Tybjaerg-Hansen A (2010) PCSK9 R46L, low-density lipoprotein cholesterol levels, and risk of ischemic heart disease: 3 independent studies and meta-analyses. J Am Coll Cardiol 55:2833-2842. doi:10.1016/j.jacc.2010.02.044
7. Kent ST, Rosenson RS, Avery CL, Chen YDI, Correa A, Cummings SR, et al. (2017) PCSK9 loss-of-function variants, low-density lipoprotein cholesterol, and risk of coronary heart disease and stroke: data from 9 studies of Blacks and Whites. Circ Cardiovasc Genet 10:e001632. doi:10.1161/CIRCGENETICS.116.001632
8. Chen S, Francioli LC, Goodrich JK, Collins RL, Kanai M, Wang Q, et al. (2022) A genome-wide mutational constraint map quantified from variation in 76,156 human genomes. bioRxiv. doi:10.1101/2022.03.20.485034
9. Stenson PD, Ball EV, Mort M, Phillips AD, Shiel JA, Thomas NST, et al. (2003) Human Gene Mutation Database (HGMD): 2003 update. Hum Mutat 21:577-581. doi:10.1002/humu.10212
10. Lacaze P, Riaz M, Sebra R, Hooper AJ, Pang J, Tiller J, et al. (2021) Protective lipid-lowering variants in healthy older individuals without coronary heart disease. Open Heart 8:e001710. doi:10.1136/openhrt-2021-001710
11. McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GRS, Thormann A, et al. (2016) The Ensembl Variant Effect Predictor. Genome Biol 17:122. doi:10.1186/s13059-016-0974-4
12. Paysan-Lafosse T, Blum M, Chuguransky S, Grego T, Pinto BL, Salazar GA, et al. (2023) InterPro in 2022. Nucleic Acids Res 51:D418-D427. doi:10.1093/nar/gkac993
13. Pundir S, Martin MJ, O'Donovan C (2017) UniProt protein knowledgebase. Methods Mol Biol 1558:41-55. doi:10.1007/978-1-4939-6783-4_2
14. Ng PC, Henikoff S (2003) SIFT: predicting amino acid changes that affect protein function. Nucleic Acids Res 31:3812-3814. doi:10.1093/nar/gkg509
15. Hecht M, Bromberg Y, Rost B (2015) Better prediction of functional effects for sequence variants. BMC Genomics 16:S1. doi:10.1186/1471-2164-16-S8-S1
16. Tavtigian SV, Deffenbaugh AM, Yin L, Judkins T, Scholl T, Samollow PB, et al. (2006) Comprehensive statistical study of 452 BRCA1 missense substitutions with classification of eight recurrent substitutions as neutral. J Med Genet 43:295-305. doi:10.1136/jmg.2005.033878
17. Capriotti E, Fariselli P, Casadio R (2005) I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure. Nucleic Acids Res 33:W306-W310. doi:10.1093/nar/gki375
18. Montanucci L, Capriotti E, Frank Y, Ben-Tal N, Fariselli P (2019) DDGun: an untrained method for the prediction of protein stability changes upon single and multiple point variations. BMC Bioinformatics 20:335. doi:10.1186/s12859-019-2923-1
19. Pires DEV, Ascher DB, Blundell TL (2014) DUET: a server for predicting effects of mutations on protein stability using an integrated computational approach. Nucleic Acids Res 42:W314-W319. doi:10.1093/nar/gku411
20. Manfredi M, Savojardo C, Martelli PL, Casadio R (2021) DeepREx-WS: a web server for characterising protein–solvent interaction starting from sequence. Comput Struct Biotechnol J 19:5791-5799. doi:10.1016/j.csbj.2021.10.016
21. Webb B, Sali A (2021) Protein structure modeling with MODEL-

LER. Methods Mol Biol 2199:239-255. doi:10.1007/978-1-0716-0892-0_14
22. Colovos C, Yeates TO (1993) Verification of protein structures: patterns of nonbonded atomic interactions. Protein Sci 2:1511-1519. doi:10.1002/pro.5560020916
23. Laskowski RA (2009) PDBsum new things. Nucleic Acids Res 37:D355-D359. doi:10.1093/nar/gkn860
24. Wiederstein M, Sippl MJ (2007) ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res 35:W407-W410. doi:10.1093/nar/gkm290
25. Johansson MU, Zoete V, Michielin O, Guex N (2012) Defining and searching for structural motifs using DeepView/Swiss-PdbViewer. BMC Bioinformatics 13:173. doi:10.1186/1471-2105-13-173
26. Burley SK, Bhikadiya C, Bi C, Bittrich S, Chao H, Chen L, et al. (2022) RCSB Protein Data Bank: tools for visualizing and understanding biological macromolecules in 3D. Protein Sci 31:e4482. doi:10.1002/pro.4482
27. Weng G, Wang E, Wang Z, Liu H, Zhu F, Li D, et al. (2019) HawkDock: a web server to predict and analyze the protein–protein complex based on computational docking and MM/GBSA. Nucleic Acids Res 47:W322-W330. doi:10.1093/nar/gkz397
28. Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Crystallogr 26: doi:10.1107/S0021889892009944
29. Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, et al. (2015) GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX 1-2:19-25. doi:10.1016/j.softx.2015.06.001
30. Abifadel M, Varret M, Rabès JP, Allard D, Ouguerram K, Devillers M, et al. (2003) Mutations in PCSK9 cause autosomal dominant hypercholesterolemia. Nat Genet 34:154-156. doi:10.1038/ng1161
31. González-Lleó AM, Sánchez-Hernández RM, Plana N, et al. (2024) Impact of PCSK9 inhibitors in glycaemic control and new-onset diabetes. Cardiovasc Diabetol 23:4. doi:10.1186/s12933-023-02077-y
32. Cohen JC, Boerwinkle E, Mosley TH Jr, Hobbs HH (2006) Sequence variations in PCSK9, low LDL, and protection against coronary heart disease. N Engl J Med 354:1264-1272. doi:10.1056/NEJMoa054013
33. Horton JD, Cohen JC, Hobbs HH (2007) Molecular biology of PCSK9: its role in LDL metabolism. Trends Biochem Sci 32:71-77. doi:10.1016/j.tibs.2006.12.008
34. Brousseau ME et al (2022) Identification of a PCSK9-LDLR disruptor peptide with in vivo function. Cell Chem Biol 29:249-258. doi:10.1016/j.chembiol.2021.08.012
35. Nurmohamed NS, Reeskamp LF (2025) Low-density lipoprotein cholesterol and coronary artery disease across ancestries—equal direction, different magnitude? N Engl J Med Evid 4:EVIDe2500256. doi:10.1056/EVIDe2500256
36. Yachida J, Fujihara K, Harada-Yamada M, Kobayashi A, Khin LM, Takizawa H, et al. (2024) coronary artery disease risk prediction by combined stratification of LDL and HDL cholesterol across different glucose statuses. Diabetes Obes Metab 26:5845-5856. doi:10.1111/dom.15956
37. Sabatine MS, Giugliano RP, Keech AC, Honarpour N, Wiviott SD, Murphy SA, et al. (2017) Evolocumab and clinical outcomes in patients with cardiovascular disease. N Engl J Med 376:1713-1722. doi:10.1056/NEJMoa1615664
38. Schwartz GG, Steg PG, Szarek M, Bhatt DL, Bittner VA, Diaz R, et al. (2018) Alirocumab and cardiovascular outcomes after acute coronary syndrome. N Engl J Med 379:2097-2107. doi:10.1056/NEJMoa1801174
39. Zhang DW, Lagace TA, Garuti R, Zhao Z, McDonald M, Horton JD, et al. (2007) Binding of proprotein convertase subtilisin/kexin type 9 to epidermal growth factor-like repeat A of low-density lipoprotein receptor decreases receptor recycling and increases degradation. J Biol Chem 282:18602-18612. doi:10.1074/jbc.M702027200
40. Piper DE, Jackson S, Liu Q, Romanow WG, Shetterly S, Thibault ST, et al. (2007) The crystal structure of PCSK9: a regulator of plasma LDL-cholesterol. Structure 15:545-552. doi:10.1016/j.str.2007.04.004
41. Ray KK, Wright RS, Kallend D, Koenig W, Leiter LA, Raal FJ, et al. (2020) Two phase 3 trials of inclisiran in patients with elevated LDL cholesterol. N Engl J Med 382:1507-1519. doi:10.1056/NEJMoa1912387
42. Lambert G, Sjouke B, Choque B, Kastelein JJP, Hovingh GK (2012) The PCSK9 decade. J Lipid Res 53:2515-2524. doi:10.1194/jlr.R026658
43. Sobati S, Shakouri A, Edalati M, Mohammadnejad D, Parvan R, Masoumi J, et al. (2020) PCSK9: a key target for the treatment of cardiovascular disease (CVD). Adv Pharm Bull 10:502-511. doi:10.34172/apb.2020.062