



Correspondence:

✉ Ayaz Ahmad, ahdayazb5@awkum.edu.pk;
Naveed Khan, naveedkhan@awkum.edu.pk

Received

04-08-25

Accepted

06-09-2025

Authors' Contributions

Concept: NH, WG, YI; Design: NH, WG, YI, AA;
Data Collection: NH, WG; Analysis: NH, WG;
Drafting: NH, WG, AA, NK

* These authors contributed equally

Copyrights

© 2025 Authors. This is an open-access article
distributed under the terms of the Creative
Commons Attribution 4.0 International License (CC
BY 4.0).



Declarations

No funding was received for this study. The authors
declare no conflict of interest. The study received
ethical approval. All participants provided informed
consent.

[“Click to Cite”](#)

Molecular Docking and Functional Analysis of APOA5 (G185C), PCSK9 (R46L, R93C), LPL (N318S), and LIPA (T16P) Genes Mutations Associated with Coronary Artery Disease

Nazia Hadi^{1*}, Wagma Gul^{1*}, Yasir Ali¹, Ayaz Ahmad¹, Naveed Khan¹

¹ Department of Biotechnology, Abdul Wali Khan University Mardan, KPK, Pakistan

ABSTRACT

Background: Coronary artery disease (CAD) is strongly influenced by genetic factors, particularly non-synonymous single nucleotide polymorphisms (nsSNPs) that alter protein structure and function. Variants within lipid metabolism-related genes such as APOA5, PCSK9, LPL, and LIPA are implicated in atherosclerosis progression, yet their molecular consequences remain incompletely defined. **Objective:** This study aimed to comprehensively characterize the structural and functional impact of selected CAD-associated nsSNPs using an integrative computational approach. **Methods:** Reported nsSNPs from the GWAS catalog were retrieved, and detailed variant data were obtained from UniProt and NCBI. Functional impacts were predicted using sequence homology-based (SIFT, PROVEAN, Mutation Assessor), machine learning-based (SNAP2, SuSPect, PolyPhen-2), and consensus predictors (Meta-SNP). Structural stability was assessed by I-Mutant, MUpro, mCSM, and DynaMut2, while evolutionary conservation, surface accessibility, and post-translational modifications were analyzed with ConSurf, NetSurf-2.0, and MusiteDeep. Protein-protein interactions were mapped via STRING, and molecular docking was performed using ClusPro and SwissDock. **Results:** APOA5 G185C, PCSK9 R46L and R93C, LPL N318S, and LIPA T16P were consistently predicted to be deleterious, with most variants exhibiting negative $\Delta\Delta G$ values indicative of destabilization. Docking analysis revealed reduced binding affinities and altered interaction residues, suggesting disruption of lipid regulatory pathways. **Conclusion:** This integrative in-silico analysis highlights critical CAD-related nsSNPs that destabilize protein structure and impair molecular interactions, underscoring their potential as biomarkers and therapeutic targets.

Keywords

Coronary artery disease, nsSNPs, APOA5, PCSK9, LPL, LIPA, protein stability, molecular docking, bioinformatics

INTRODUCTION

Coronary artery disease (CAD), also referred to as ischemic heart disease (IHD) or coronary heart disease (CHD), remains a leading cause of mortality in industrialized nations. Despite significant advancements in both prevention and treatment (1), CAD continues to pose a major global health challenge, accounting for more than 3.9 million deaths in Europe and 1.8 million in the European Union annually. In the United States, more than 18.2 million individuals are affected, with approximately 805,000 developing acute coronary syndrome (ACS) each year (2). By 2030, the global mortality associated with CHD is projected to reach nearly 9.245 million, highlighting the increasing global burden reported by the World Health Organization (3).

The underlying cause of CAD is primarily the buildup of atherosclerotic plaques in the epicardial coronary arteries, which restricts blood flow (4). Although the precise pathological and physiological mechanisms remain incompletely understood, both genetic and environmental factors play critical roles in disease progression. Established risk factors include advanced age, dyslipidemia, obesity, hypertension, diabetes mellitus, tobacco use, alcohol consumption, and unhealthy lifestyle practices. Genome-wide association studies (GWAS) have identified over 230 genetic variants significantly associated with CAD (5). Importantly, research suggests that the pathophysiology of CAD extends beyond epicardial atherosclerotic plaques, with coronary microcirculation also playing a pivotal role in disease manifestation (6).

CAD presents with a wide clinical spectrum, ranging from asymptomatic subclinical atherosclerosis to severe outcomes such as angina pectoris, acute myocardial infarction (MI), and sudden cardiac death (SCD) (7). This broad phenotypic variability results from complex interactions between genetic mutations and environmental exposures, including dietary patterns, physical activity, tobacco use, and comorbidities (8). Among the genes implicated in lipoprotein metabolism, APOA5, PCSK9, LPL, and LIPA are of particular importance. Since lipoprotein metabolism is central to atherosclerotic plaque formation, mutations in these genes are strongly linked with CAD.

The lipoprotein lipase (LPL) gene, located on chromosome 8p21.3, encodes an enzyme critical for lipid metabolism by mediating the hydrolysis of very low-density lipoproteins (VLDL) into low-density lipoproteins (LDL-C). Mutations in LPL are associated with lipoprotein lipase deficiency, a rare autosomal recessive disorder characterized by markedly elevated triglyceride levels, lactescent serum, reduced concentrations of

HDL-C and LDL-C, and clinical features such as eruptive xanthomas, abdominal pain, hepatosplenomegaly, and, in some cases, early-onset atherosclerotic CAD (9).

The proprotein convertase subtilisin/kexin type 9 (PCSK9) gene, located on chromosome 1p32.3, encodes a protease that downregulates LDL receptor (LDLR) expression in both hepatic and extrahepatic tissues (10). Mutations in PCSK9 reduce LDL receptor density, leading to elevated plasma total cholesterol and LDL-C levels. Clinically, such alterations are associated with tendon xanthomas, premature CAD, myocardial infarction, and ischemic stroke (11).

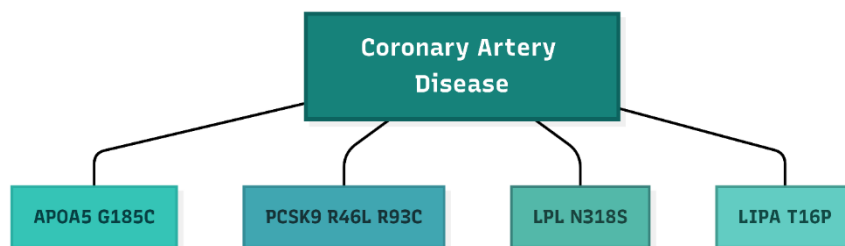


Figure 1 Schematic Representation of CAD-Associated Gene Variants

The APOA5 gene encodes apolipoprotein A-V (apoA-V), a key regulator of plasma triglyceride concentrations. Genetic variants in APOA5 have been consistently associated with altered lipid metabolism and an increased risk of CAD (12).

The LIPA gene encodes lysosomal acid lipase (LAL), an enzyme essential for hydrolyzing cholesterol esters and triglycerides derived from internalized lipoproteins. Mutations in LIPA have been strongly linked with an elevated risk of CAD (13).

MATERIAL AND METHODS

Data Collection

In this study, we evaluated single-nucleotide polymorphisms (SNPs) reported in genome-wide association studies (GWAS) that are associated with coronary artery disease (CAD). Four key genes involved in lipid metabolism were selected for detailed analysis. Protein sequences corresponding to these genes were retrieved from the UniProt database. Mutation-related information, including allele variations, amino acid changes, chromosomal locations, and global minor allele frequencies, was extracted from the NCBI database.

Prediction of the Damaging Impact of Reported nsSNPs

To evaluate the potential damaging effects of non-synonymous SNPs (nsSNPs) on protein structure and function, eleven *in silico* prediction tools were employed. These tools were grouped into two categories: (i) predictors of functional impact (SIFT, PROVEAN, Mutation Assessor, CADD (14), PolyPhen-2, and SNAP2), and (ii) predictors of pathogenicity (PhD-SNP and SuSPect). Additionally, consensus-based methods, such as Meta-SNP, were included to enhance predictive reliability.

The selected tools represent four major computational approaches: sequence homology-based, supervised machine learning-based, protein sequence/structure-based, and consensus-based prediction. Specifically, SIFT (15), PROVEAN, and Mutation Assessor (16) were used for sequence homology-driven analysis of functional consequences. SNAP2, based on a neural network (17), and SuSPect, based on support vector machines (18), provided machine learning predictions. PolyPhen-2, which integrates sequence, phylogenetic, and structural features, was included due to its balanced sensitivity and specificity (19). Meta-SNP (20) was employed as a consensus approach that integrates outputs from PANTHER, PhD-SNP, SIFT, and SNAP, thereby increasing overall confidence in predictions.

Structural Stability and Dynamic Flexibility Analysis of Missense Variants

I-Mutant 2.0: Protein stability changes induced by missense SNPs were predicted using the I-Mutant 2.0 web server. This tool employs a Support Vector Machine (SVM)-based algorithm to estimate changes in Gibbs free energy ($\Delta\Delta G$). A positive $\Delta\Delta G$ (>0 kcal/mol) indicates increased stability, whereas a negative $\Delta\Delta G$ (<0 kcal/mol) indicates decreased stability (21).

MUpro: MUpro, which uses SVMs and neural networks, was used to further assess protein stability. Scores below 0 suggest destabilization, while scores above 0 indicate increased stability (22).

MEDUSA: Protein flexibility was predicted using the MEDUSA web server (23). The tool classifies amino acid residues into two, three, or five flexibility classes based on their evolutionary and physicochemical properties. Input amino acid sequences in FASTA format were analyzed, and flexibility classes were assigned accordingly.

Analysis of Evolutionary Conservation, Surface Accessibility, and PTMs

ConSurf: Conservation of amino acid residues was analyzed using the ConSurf server (24), which estimates evolutionary conservation based on phylogenetic relationships among homologous sequences. This approach helps distinguish true conservation from limited evolutionary divergence.

NetSurf-2.0: Surface accessibility, solvent exposure, and structural disorder were predicted using NetSurf-2.0 (25). This neural network-based tool processes primary protein sequences to predict solvent accessibility and secondary structure.

MusiteDeep: Post-translational modification (PTM) sites were predicted using MusiteDeep (26), a deep learning-based server capable of directly analyzing raw protein sequences for multiple PTM types.

Structural Impact Assessment of Mutations

DynaMut2: To evaluate the effect of nsSNPs on protein dynamics and flexibility, DynaMut2 was used (27). Wild-type protein structures (PDB format) were compared against mutant models to predict alterations in stability and interactions.

mCSM: The destabilizing effects of mutations were further examined using mCSM (28), a machine learning-based method that incorporates structural features to estimate changes in folding free energy ($\Delta\Delta G$), classifying mutations as stabilizing or destabilizing.

3D Modeling and Structural Analysis

Wild-type and mutant protein structures were modeled using SWISS-MODEL (29). The process involved template identification, alignment, model building, and quality evaluation. Automatic mode employed BLAST for template selection. Generated models were validated using SAVES 6.0 (30). Structural comparison between wild-type and mutant proteins was performed using TM-align (31). TM-score was calculated to assess fold similarity, while RMSD values provided insights into structural deviations, with higher RMSD values reflecting greater structural divergence.

Functional Protein Association Networks

Protein–protein interaction networks were constructed using the STRING database (32). STRING integrates ortholog clustering and diverse interaction datasets to identify potential functional associations. In this study, STRING was employed to uncover CAD-related protein interactions and provide biological insights into their disease relevance.

Gene–Drug Interaction Analysis

Drug–gene associations were explored using a drug–gene interaction database (33). This platform integrates therapeutic connection data to identify pharmacological agents with potential targeting effects on CAD-associated genes, enabling candidate drug prioritization.

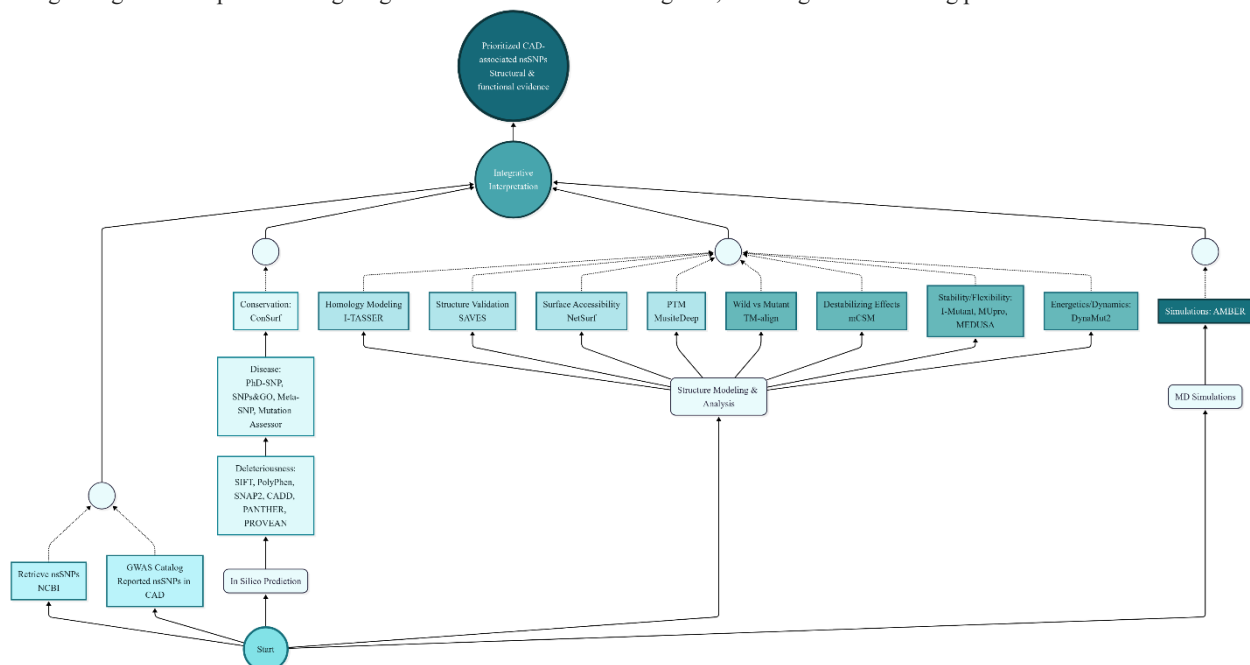


Figure 2 Workflow of integrative in-silico analysis of CAD-associated nsSNPs

Computational Docking Analysis of Protein–Protein Interactions

Protein–protein docking simulations were conducted using the ClusPro server (34). High-confidence interactors identified from STRING were used as ligands, while CAD-associated proteins served as receptors. Protein structures were obtained from the PDB or modeled with SWISS-MODEL. Docking simulations were performed using default ClusPro parameters, including free energy scoring and rotational sampling at 180° intervals.

Protein–Ligand Docking Analysis

Protein–ligand docking was conducted using SWISSdock (35). Ligands were retrieved from DrugBank in MOL2 format, while protein structures were provided in PDB format. A protein–ligand complex was modeled to estimate binding affinity and interaction profiles. Post-docking analysis was carried out using BIOVIA Discovery Studio to visualize binding conformations and interaction patterns.

RESULTS

A total of five nsSNPs across APOA5, PCSK9, LPL, and LIPA genes were prioritized for analysis based on their reported association with coronary artery disease (CAD) (Table 1). These included APOA5 (G185C), PCSK9 (R46L and R93C), LPL (N318S), and LIPA (T16P), each mapped to distinct chromosomal positions with minor allele frequencies ranging from 0.0026 to 0.2861.

Pathogenicity predictions from multiple in silico tools provided consistent evidence of deleterious effects for most variants (Table 2). APOA5 G185C and LIPA T16P were strongly classified as disease-associated by SNPs&GO, PhD-SNP, and Meta-SNP, with corroborating evidence from

PolyPhen-2 and PROVEAN. For PCSK9, both R46L and R93C were predicted deleterious, though R46L showed lower damaging scores in SIFT and PANTHER. LPL N318S was largely classified as neutral, with mixed predictions across tools, suggesting possible context-dependent effects. Protein stability assessments further supported the destabilizing nature of these variants (Table 3). APOA5 G185C, PCSK9 R46L, and LIPA T16P consistently showed negative $\Delta\Delta G$ values in I-Mutant and MUpro, indicating reduced stability, while PCSK9 R93C displayed contradictory outcomes—stability increase in I-Mutant but strong destabilization in MUpro—highlighting the importance of using multiple predictive models. MEDUSA confirmed reduced structural flexibility for APOA5, PCSK9 R93C, LPL, and LIPA mutants, while PCSK9 R46L displayed a tendency toward increased flexibility. Evolutionary conservation and solvent accessibility analyses demonstrated that these variants occur at functionally relevant residues (Table 4). APOA5 G185C, PCSK9 R46L, and LIPA T16P were solvent-exposed, while PCSK9 R93C was highly conserved and exposed at 52%, suggesting functional disruption at critical sites. Disordered residue predictions revealed marked disorder for PCSK9 R46L (60%) and LIPA T16P (100%), indicating destabilization of flexible regions. Importantly, no post-translational modification sites were predicted at these positions, ruling out confounding effects.

Dynamic stability modeling using DynaMut2 (Table 5) confirmed destabilizing effects for all variants except LPL N318S, which exhibited a stabilizing $\Delta\Delta G$ of +0.42 kcal/mol. mCSM predictions (Table 6) aligned closely, with all variants classified as destabilizing, reinforcing the consensus that these nsSNPs alter protein structural integrity.

Comparative structural modeling (Figure 3) revealed that mutant proteins displayed conformational deviations relative to wild-type. APOA5 G185C showed subtle alterations in ϕ - ψ distributions; PCSK9 R46L and R93C introduced pronounced deviations in loop and helical regions; LPL N318S introduced local distortions near the catalytic site; and LIPA T16P disrupted N-terminal secondary structure. Ramachandran plots consistently confirmed increased outlier residues across mutants, reflecting compromised folding.

Protein-protein interaction (PPI) network analysis (Figure 4) provided further insights. APOA5 interacted strongly with other apolipoproteins, PCSK9 with LDLR and HMGCR, LPL with GPIHBP1 and APOC2, and LIPA with cholesterol metabolism enzymes. Mutations in these genes are therefore likely to propagate dysfunction through lipid regulation networks, amplifying CAD risk.

Molecular docking analysis demonstrated significant alterations in ligand binding (Table 7; Figure 5). Wild-type APOA5 exhibited strong binding with APOC3 (−1208.9 kcal/mol), which was slightly weakened in the G185C mutant. PCSK9 wild-type bound LDLR with a very strong affinity (−1911.1 kcal/mol), but both R46L and R93C mutants displayed substantially reduced binding (−1633.9 and −1536.5 kcal/mol, respectively), consistent with impaired receptor interaction. LPL N318S showed reduced binding affinity to GPIHBP1 compared to wild-type (−1396.3 vs. −1432.8 kcal/mol). Similarly, LIPA T16P significantly weakened binding to SCGB1D2 (−1013.9 vs. −1343.1 kcal/mol). These altered binding profiles support the hypothesis that nsSNPs disrupt key protein-ligand and protein-protein interactions essential for lipid metabolism.

Finally, drug-gene interaction analysis (Table 8) highlighted that several variants mapped to proteins targeted by established lipid-lowering drugs. Rosuvastatin (APOA5), lomitapide mesylate (PCSK9), clofibrate (LPL), and lovastatin (LIPA) demonstrated varying interaction scores, underscoring their therapeutic relevance in CAD.

Table 1. Detailed information of nsSNPs in selected CAD-associated genes

Gene	rsID	Chromosomal Position	Nucleotide Change (hg38)	Amino Acid Change	Global Minor Allele Frequency	Reference
APOA5	rs2075291	chr11:116790676	NC_000011.10:g.116790676C>A	G185C	A = 0.0114	29263402
PCSK9	rs11591147	chr1:55039974	NC_000001.11:g.55039974G>T	R46L	T = 0.0064	29212778, 28714975
PCSK9	rs151193009	chr1:55043912	NC_000001.11:g.55043912C>T	R93C	T = 0.0026	33020668
LPL	rs268	chr8:19956018	NC_000008.11:g.19956018A>G	N318S	G = 0.0052	36474045
LIPA	rs1051338	chr10:89247603	NC_000010.11:g.89247603T>G	T16P	G = 0.2861	36474045

Table 2. Predictions and scores of nsSNPs from bioinformatics tools

Gene	rsID	Amino Acid Change	SNPs&G O (Score)	PhD-SNP (Score)	PANTHER	SIFT	PolyPhen-2	PROVEAN (Cutoff -2.5)	SuSPect	SNAP 2	Mutation Assessor	CA DD	Meta-SNP
APOA5	rs2075291	G185C	Disease (0.594)	Disease (4)	Probably benign (0.19)	Affects function (0.03)	Probably damaging (0.997)	Deleterious (-3.444)	Neutral polymorphism (25)	Effect (0.650)	Medium (0.685)	Like ly benign (19)	Neutral (0.468)
PCSK9	rs11591147	R46L	Disease (0.649)	Disease (2)	Probably benign (0.02)	Tolerated (0.10)	Benign (0.001)	Neutral (-0.236)	Neutral polymorphism (17)	Neutral (0.485)	Low (0.202)	Like ly benign (0.028)	Disease (0.506)
PCSK9	rs151193009	R93C	Disease (0.721)	Disease (0)	Probably benign (0.02)	Tolerated (0.06)	Probably damaging (1.000)	Deleterious (-3.172)	Neutral polymorphism (31)	Effect (0.540)	Medium (0.788)	Like ly benign (24)	Disease (0.519)
LPL	rs268	N318S	Neutral (0.235)	Neutral (2)	Probably damaging (0.57)	Tolerated (0.52)	Benign (0.143)	Neutral (-1.085)	Neutral polymorphism (21)	Neutral (0.390)	Low (0.408)	Like ly benign (21)	Neutral (0.430)
LIPA	rs1051338	T16P	Disease (0.763)	Disease (0)	Probably benign (0.27)	Tolerated (0.14)	Benign (0.002)	Neutral (-1.212)	Neutral polymorphism (16)	Effect (0.550)	Neutral (0.065)	Like ly benign (11)	Disease (0.736)

Table 3. Stability and flexibility prediction of nsSNPs using I-Mutant2.0, MUpro, and MEDUSA

Gene	rsID	Amino Acid Change	I-Mutant $\Delta\Delta G$ (kcal/mol)	I-Mutant Prediction	MUpro $\Delta\Delta G$ (kcal/mol)	MUpro Prediction	MEDUSA Flexibility	MEDUSA Probability (%)	C-score
APOA5	rs2075291	G185C	-0.69	Decrease stability	-0.676	Decrease stability	Rigid	45	<0.5
PCSK9	rs11591147	R46L	-0.33	Decrease stability	-0.382	Decrease stability	Flexible	49	<0.5
PCSK9	rs151193009	R93C	+0.71	Increase stability	-0.812	Decrease stability	Rigid	40	<0.5
LPL	rs268	N318S	-0.63	Decrease stability	-1.103	Decrease stability	Rigid	44	<0.5
LIPA	rs1051338	T16P	-0.52	Decrease stability	-1.136	Decrease stability	Rigid	56	<0.5

Table 4. Conservation, solvent accessibility, and post-translational modification (PTM) analysis of selected nsSNPs

Gene	rsID	Amino Acid Change	ConSurf Score (Conservation)	Relative Surface Accessibility (%)	Disordered Residues (%)	PTM Prediction (MusiteDeep)
APOA5	rs2075291	G185C	Variable (3)	34 (Exposed)	0	No PTM predicted
PCSK9	rs11591147	R46L	Variable (3)	46 (Exposed)	60	No PTM predicted
PCSK9	rs151193009	R93C	Conserved (7)	52 (Exposed)	0	No PTM predicted
LPL	rs268	N318S	Average (5)	45 (Exposed)	1	No PTM predicted
LIPA	rs1051338	T16P	Variable (2)	84 (Exposed)	100	No PTM predicted

Table 5. Predicted stability changes of nsSNPs using DynaMut2

Gene	rsID	Amino Acid Change	$\Delta\Delta G$ Stability (kcal/mol)	Prediction
APOA5	rs2075291	G185C	-0.39	Destabilizing
PCSK9	rs11591147	R46L	-0.23	Destabilizing
PCSK9	rs151193009	R93C	-0.31	Destabilizing
LPL	rs268	N318S	+0.42	Stabilizing
LIPA	rs1051338	T16P	-0.02	Destabilizing

Table 6. Predicted alterations in protein stability upon nsSNPs using mCSM

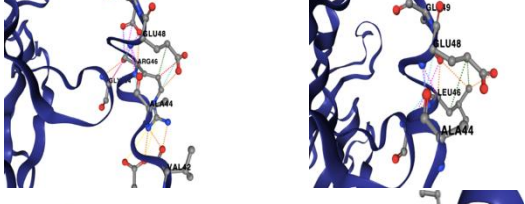
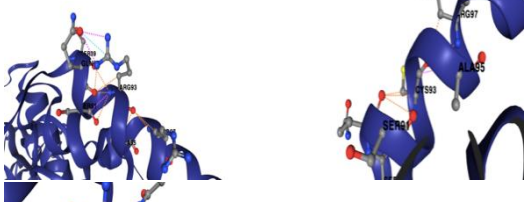
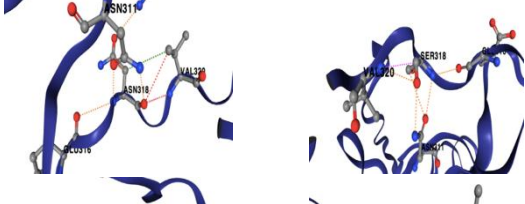
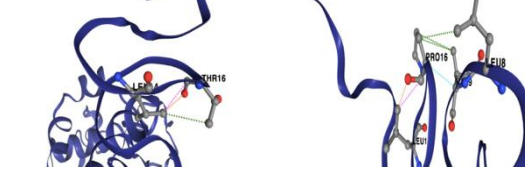
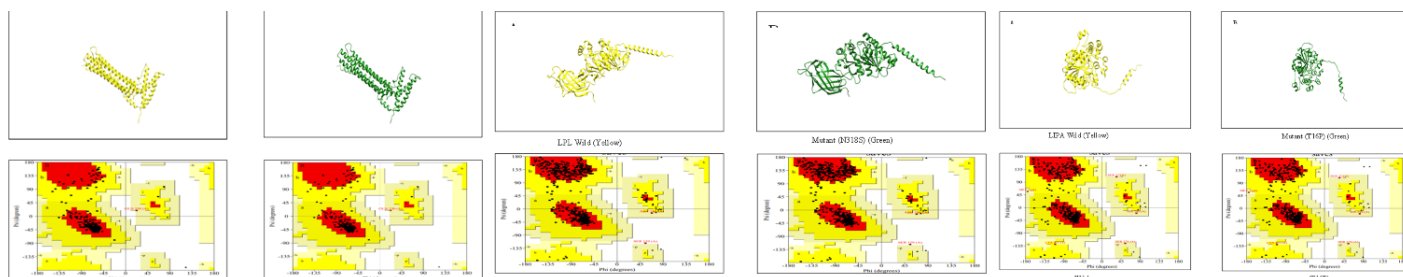
Gene	rsID	Amino Acid Change	$\Delta\Delta G$ (kcal/mol)	Stability Prediction	Structural Representation
APOA5	rs2075291	G185C	+0.39	Destabilizing	
PCSK9	rs11591147	R46L	+0.23	Destabilizing	
PCSK9	rs151193009	R93C	+0.23	Destabilizing	
LPL	rs268	N318S	+0.31	Destabilizing	
LIPA	rs1051338	T16P	+0.02	Destabilizing	

Table 7. Molecular docking analysis of wild-type and mutant CAD-associated proteins with their interacting ligands

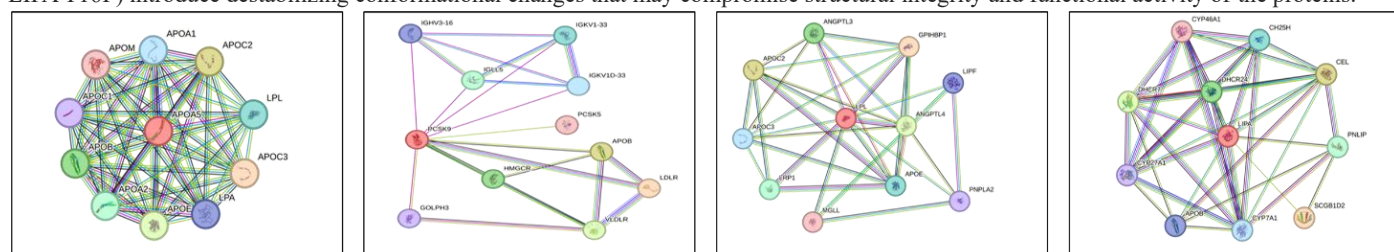
Protein Variant	Ligand	Binding Energy (kcal/mol)	Key Interacting Residues
APOA5 (WT)	APOC3	-1208.9	LEU14, PHE18, ARG24, HIS45, GLU138, GLN145, GLN148, GLU149, TYR194, ARG204, ARG211, ARG282, GLN286
APOA5 (G185C)	APOC3	-1205.2	LEU14, ARG24, HIS45, GLU138, GLN145, GLN148, GLU149, TYR194, ARG204, ARG211, ARG282, GLN286
PCSK9 (WT)	LDLR	-1911.1	SER6, ARG7, ARG8, SER9, TRP10, TRP11, PRO12, LEU17, LEU18, LEU19, ALA26, GLY27, ASP37, SER47, GLU48, GLU49, ARG46, VAL42, GLU40, HIS87, GLN90, ARG93, ARG96, ARG97, ARG104, TYR293, ARG306, ARG476, GLU567
PCSK9 (R46L)	LDLR	-1633.9	PHE64, HIS65, ARG66, HIS87, SER89, GLU92, ARG104, LYS136, THR214, ARG215, HIS217, ASP238, ARG272, TYR293, ARG306, GLU332, ARG357, ASP374, THR377, CYS378, ARG476
PCSK9 (R93C)	LDLR	-1536.5	LEU17, ALA26, SER46, HIS66, ARG68, HIS87, SER90, SER93, HIS217, TYR293, ARG306, THR377, ASP374, ARG476, GLU566
LPL (WT)	GPIHBP1	-1432.8	ARG290, SER292, LEU303, SER304, ARG306, LYS307, ARG309, LYS327, LYS346, HIS348, GLY351, SER354, THR385, ASN386, TRP420, ARG432, LYS434
LPL (N318S)	GPIHBP1	-1396.3	ARG219, ARG290, SER292, LEU303, SER304, ARG306, LYS307, ARG309, ARG324, LYS327, LYS346, GLY351, THR352, SER354, THR385, ASN386, TRP420, ARG432, LYS434
LIPA (WT)	SCGB1D2	-1343.1	THR27, HIS250, ARG270, SER275, HIS344, LEU381
LIPA (T16P)	SCGB1D2	-1013.9	THR27, HIS250, ARG270, SER275, HIS344, LEU381

Table 8. Predicted drug-gene interactions of CAD-associated variants with approved antihypercholesterolemic agents

Gene	Drug	Indication	Interaction Score	DrugBank ID
APOA5	Rosuvastatin	Antihypercholesterolemic agent	0.87	DB01098
PCSK9	Lomitapide Mesylate	Antihypercholesterolemic agent	0.80	DB08827
LPL	Clofibrate	Anticholesteremic agent	0.11	DB00636
LIPA	Lovastatin	Antihypercholesterolemic agent	0.44	DB00227

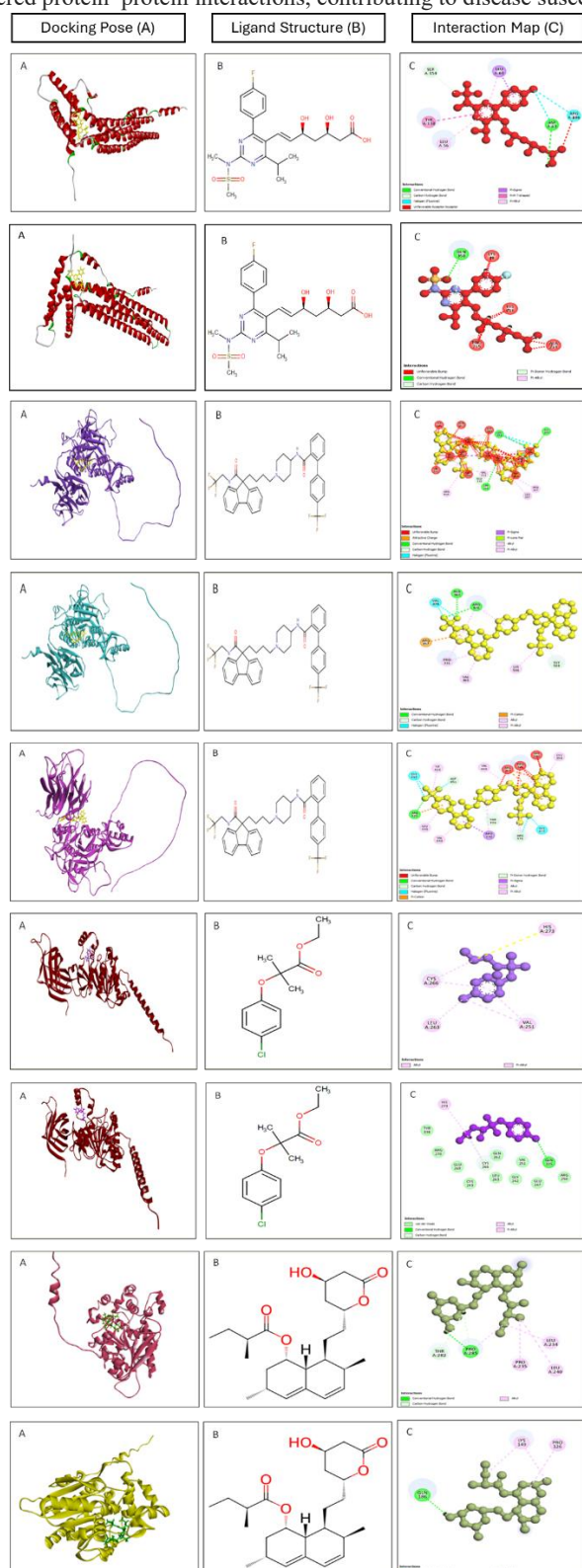
**Figure 3 Comparative structural modeling and Ramachandran plot validation of wild-type and mutant CAD-associated proteins.**

This figure presents the modeled three-dimensional (3D) structures of coronary artery disease (CAD)-associated proteins in both wild-type and mutant forms, alongside their respective Ramachandran plots for conformational validation. In panel A, APOA5 wild-type (yellow) and mutant G185C (green) are shown, with the Ramachandran plots highlighting subtle deviations in ϕ - ψ angle distribution that reflect destabilizing effects introduced by the mutation. Panel B displays PCSK9 wild-type (yellow) and mutants R46L (green) and R93C (blue), where conformational differences are evident in loop and helical regions, and Ramachandran plots reveal altered residue distributions in disallowed regions, consistent with reduced stability. Panel C illustrates LPL wild-type (yellow) and mutant N318S (green), where structural superimposition reveals local distortions near the catalytic domain, supported by Ramachandran plots showing increased outlier residues indicative of impaired folding. Panel D shows LIPA wild-type (yellow) and mutant T16P (green), where the mutation introduces perturbations in the N-terminal secondary structure, with Ramachandran plots confirming significant deviations in residue orientation that reduce stability and flexibility. Collectively, these comparative modeling and validation analyses confirm that CAD-related nsSNPs (APOA5 G185C, PCSK9 R46L and R93C, LPL N318S, and LIPA T16P) introduce destabilizing conformational changes that may compromise structural integrity and functional activity of the proteins.

**Figure 4 Protein-protein interaction (PPI) networks of CAD-associated genes predicted by STRING.**

The interaction maps illustrate the functional associations of (A) APOA5, (B) PCSK9, (C) LPL, and (D) LIPA with their respective interacting partners. Each node represents a protein, and edges indicate predicted interactions supported by experimental data, co-expression, text mining, and curated databases. APOA5 showed extensive connectivity with other apolipoproteins (APOA1, APOC1, APOC2, APOC3, APOB), highlighting its central role in lipid transport and metabolism. PCSK9 was strongly linked with LDLR, APOB, and HMGCR, reflecting its involvement in

cholesterol regulation. LPL demonstrated multiple connections with GPIHBP1, APOC2, and APOE, consistent with its role in triglyceride hydrolysis and lipid clearance. LIPA clustered with cholesterol metabolism proteins (CYP27A1, CYP46A1, DHCR24), confirming its role in lysosomal lipid degradation. Overall, these networks emphasize that mutations in CAD-associated genes may disrupt critical lipid regulatory pathways through altered protein–protein interactions, contributing to disease susceptibility.



This figure presents the molecular docking analysis of wild-type and mutant CAD-associated proteins with their respective ligands, arranged in structured multilayered panels. Each row corresponds to a specific protein–ligand interaction, with three aligned components: Panel A (Docking Pose) shows the 3D representation of the ligand bound within the active site of the protein, highlighting conformational fit and spatial orientation; Panel B (Ligand Structure) depicts the 2D chemical structure of the corresponding ligand for clarity; and Panel C (Interaction Map) provides detailed residue-level visualization of non-covalent interactions, including hydrogen bonds, hydrophobic contacts, π – π stacking, and electrostatic interactions. Together, these panels illustrate how nsSNP-induced structural changes influence docking stability and binding conformations. Specifically, APOA5 with Rosuvastatin, PCSK9 with Lomitapide Mesylate, LPL with Clofibrate, and LIPA with Lovastatin show variations in binding affinities and interacting residues between wild-type and mutant proteins, reflecting altered drug–target interactions. This integrated depiction underscores the mechanistic impact of CAD-related mutations on protein–ligand dynamics, offering insights into therapeutic potential and druggability of the studied variants.

Figure 5 Comparative docking analysis of wild-type and mutant CAD-associated proteins with therapeutic ligands.

DISCUSSION

Single nucleotide polymorphisms (SNPs) are the most prevalent form of genetic variation, occurring across both coding and non-coding regions of the human genome. Among these, non-synonymous SNPs (nsSNPs) are of particular clinical relevance, as they result in amino acid substitutions that can disrupt protein stability, folding, and intermolecular interactions. Such structural perturbations can compromise essential biological processes and contribute to the etiology of complex diseases, including coronary artery disease (CAD) (1–3). The characterization of nsSNPs in CAD-related genes is therefore critical for advancing biomarker discovery and therapeutic development.

In this study, we applied a multi-layered in-silico framework to systematically evaluate five nsSNPs within four CAD-associated genes—APOA5, PCSK9, LPL, and LIPA. To overcome the inherent limitations of single predictors, we incorporated a range of computational approaches spanning evolutionary conservation-based tools (SIFT, PROVEAN, Mutation Assessor) (14–16), machine learning-based classifiers (PolyPhen-2, SNAP2, SuSPect) (17–19), and consensus methods such as Meta-SNP (20). Despite methodological variability, APOA5 G185C, PCSK9 R46L and R93C, and LIPA T16P were consistently predicted as deleterious, while LPL N318S exhibited a mixed profile, being largely neutral in some algorithms but pathogenic in others. This highlights the value of integrative predictions in distinguishing high-confidence risk variants from background polymorphisms.

Structural and stability analyses further reinforced these findings. Solvent accessibility assessments revealed that most variants occurred at surface-exposed residues, predisposing them to disrupt ligand recognition or protein–protein interactions (25,26). Stability prediction tools, including I-Mutant, MUpro, mCSM, and DynaMut2, demonstrated predominantly negative $\Delta\Delta G$ values (21,22,27,28), confirming destabilizing effects. Such decreases in thermodynamic stability are hallmarks of disease-associated mutations and may impair proper protein folding, leading to functional insufficiency. Interestingly, LPL N318S displayed a stabilizing effect in DynaMut2, suggesting a context-dependent mechanism that warrants further biochemical validation.

At the systems level, protein–protein interaction (PPI) mapping revealed that these variants perturb key regulatory hubs involved in lipid transport and cholesterol metabolism. For example, APOA5 clustered with multiple apolipoproteins, PCSK9 maintained critical associations with LDLR and HMGCR, and LIPA was linked to cholesterol degradation enzymes (32). Disruption of such densely interconnected networks may impair lipoprotein clearance, accelerate atherosclerosis, and promote CAD progression.

Molecular docking analyses provided further mechanistic insights. Mutations in PCSK9 substantially weakened its binding affinity with LDLR, consistent with its known role in cholesterol regulation (10,11). Similarly, LPL N318S and LIPA T16P displayed reduced interaction energies with GPIHBP1 and SCGB1D2, respectively, supporting their potential pathogenicity in lipid metabolism (9,13). APOA5 G185C maintained near-native binding energies but altered specific residue interactions with APOC3, suggesting subtle functional consequences. These results collectively indicate that nsSNP-induced perturbations compromise not only structural stability but also intermolecular communication, thereby amplifying their pathogenic impact.

Taken together, the findings emphasize the utility of integrated computational pipelines in dissecting the molecular consequences of nsSNPs. By combining sequence-based prediction, conservation analysis, structural modeling, stability assessments, PPI network mapping, and molecular docking, this study provides a robust framework for prioritizing variants with the greatest likelihood of functional impact in CAD.

CONCLUSION

This study comprehensively evaluated five nsSNPs across four CAD-associated genes—APOA5, PCSK9, LPL, and LIPA—using an integrative in-silico strategy. Pathogenicity predictors consistently identified APOA5 G185C, PCSK9 R46L and R93C, and LIPA T16P as deleterious (14–20), while structural and stability models revealed destabilizing conformational changes in most variants (21,22,27,28). Protein–protein interaction analysis demonstrated that these genes act as central regulators in lipid metabolic pathways (32), and molecular docking confirmed reduced binding affinities and altered interaction profiles with physiological partners (9–13,34,35).

The strength of this work lies in its multi-dimensional approach, which bridges molecular predictions with network-level consequences, thereby generating a comprehensive molecular portrait of CAD-related genetic variants. Such in-silico frameworks are valuable for identifying high-risk mutations, informing drug–gene interactions, and guiding experimental priorities.

Nevertheless, computational predictions remain probabilistic and require validation. Future studies should integrate biochemical assays, cell-based functional models, and patient-derived data to confirm the pathogenicity of these variants and to explore their potential as diagnostic markers or therapeutic targets. Ultimately, the synergy of computational and experimental approaches will be essential for advancing precision medicine strategies in coronary artery disease.

REFERENCES

- Wilkins E, Wilson L, Wickramasinghe K, Bhatnagar P, Leal J, Luengo-Fernandez R, et al. European cardiovascular disease statistics 2017. 2017.
- Muse ED, Chen S-F, Torkamani AJCcr. Monogenic and polygenic models of coronary artery disease. 2021;23:1-12.
- Nordlie MA, Wold LE, Kloner RAJJom, cardiology c. Genetic contributors toward increased risk for ischemic heart disease. 2005;39(4):667-79.
- Mach F, Baigent C, Catapano AL, Koskinas KC, Casula M, Badimon L, et al. 2019 ESC/EAS Guidelines for the management of dyslipidaemias: lipid modification to reduce cardiovascular risk: The Task Force for the management of dyslipidaemias of the European Society of Cardiology (ESC) and European Atherosclerosis Society (EAS). 2020;41(1):111-88.
- Consortium CD, Deloukas P, Kanoni S, Willenborg C, Farrall M, Assimes TL, et al. Large-scale association analysis identifies new risk loci for coronary artery disease. 2013;45(1):25-33.
- Visseren FL, Mach F, Smulders YM, Carballo D, Koskinas KC, Bäck M, et al. 2021 ESC Guidelines on cardiovascular disease prevention in clinical practice: Developed by the Task Force for cardiovascular disease prevention in clinical practice with representatives of the European Society of Cardiology and 12 medical societies With the special contribution of the European Association of Preventive Cardiology (EAPC). 2021;42(34):3227-337.
- Dai X, Wiernek S, Evans JP, Runge MSJWjoc. Genetics of coronary artery disease and myocardial infarction. 2016;8(1):1.
- Whelton PK, Carey RM, Aronow WSJII. Acc/aha/aapa/abc/acpm/ags/APhA/ASH/ASPC/nma/pcna guideline for the prevention, Detection, evaluation, and management of high blood pressure in adults: a Report of the American College of Cardiology/American heart Association. Task force on clinical practice guidelines//J. Am. Coll. Cardiol.-2017.-Nov 13. 2018;7(1):68-74.
- Gehrisch SJCar. Common mutations of the lipoprotein lipase gene and their clinical significance. 1999;1(1):70-8.
- Schmidt RJ, Beyer TP, Bensch WR, Qian Y-W, Lin A, Kowala M, et al. Secreted proprotein convertase subtilisin/kexin type 9 reduces both hepatic and extrahepatic low-density lipoprotein receptors in vivo. 2008;370(4):634-40.

11. Cohen JC, Boerwinkle E, Mosley Jr TH, Hobbs HHJNEJoM. Sequence variations in PCSK9, low LDL, and protection against coronary heart disease. 2006;354(12):1264-72.
12. Dai W, Zhang Z, Yao C, Zhao SJLiH, Disease. Emerging evidences for the opposite role of apolipoprotein C3 and apolipoprotein A5 in lipid metabolism and coronary artery disease. 2019;18:1-7.
13. Evans TD, Zhang X, Clark RE, Alisio A, Song E, Zhang H, et al. Functional characterization of LIPA (lysosomal acid lipase) variants associated with coronary artery disease. 2019;39(12):2480-91.
14. Rentzsch P, Witten D, Cooper GM, Shendure J, Kircher MJNar. CADD: predicting the deleteriousness of variants throughout the human genome. 2019;47(D1):D886-D94.
15. Sim N-L, Kumar P, Hu J, Henikoff S, Schneider G, Ng PCJNar. SIFT web server: predicting effects of amino acid substitutions on proteins. 2012;40(W1):W452-W7.
16. Reva B, Antipin Y, Sander CJNar. Predicting the functional impact of protein mutations: application to cancer genomics. 2011;39(17):e118-e.
17. Hecht M. Prediction of functional effects for sequence variants: Technische Universität München; 2015.
18. Yates CM, Filippis I, Kelley LA, Sternberg MJJJomb. SuSPect: enhanced prediction of single amino acid variant (SAV) phenotype using network features. 2014;426(14):2692-701.
19. Adzhubei I, Jordan DM, Sunyaev SRJCpihg. Predicting functional effect of human missense mutations using PolyPhen-2. 2013;76(1):7.20. 1-7.. 41.
20. Capriotti E, Altman RB, Bromberg YJBg. Collective judgment predicts disease-associated single nucleotide variants. 2013;14:1-9.
21. Capriotti E, Fariselli P, Casadio RJNar. I-Mutant2. 0: predicting stability changes upon mutation from the protein sequence or structure. 2005;33(suppl_2):W306-W10.
22. Worth CL, Preissner R, Blundell TLJNar. SDM—a server for predicting effects of mutations on protein stability and malfunction. 2011;39(suppl_2):W215-W22.
23. Vander Meersche Y, Cretin G, de Brevern AG, Gelly J-C, Galochkina TJJomb. MEDUSA: prediction of protein flexibility from sequence. 2021;433(11):166882.
24. Ashkenazy H, Erez E, Martz E, Pupko T, Ben-Tal NJNar. ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. 2010;38(suppl_2):W529-W33.
25. Klausen MS, Jespersen MC, Nielsen H, Jensen KK, Jurtz VI, Soenderby CK, et al. NetSurfP-2.0: Improved prediction of protein structural features by integrated deep learning. 2019;87(6):520-7.
26. Guo L, Wang Y, Xu X, Cheng K-K, Long Y, Xu J, et al. DeepPSP: a global–local information-based deep neural network for the prediction of protein phosphorylation sites. 2020;20(1):346-56.
27. Rodrigues CH, Pires DE, Ascher DBJPS. DynaMut2: Assessing changes in stability and flexibility upon single and multiple point missense mutations. 2021;30(1):60-9.
28. Pires DE, Ascher DB, Blundell TLJB. mCSM: predicting the effects of mutations in proteins using graph-based signatures. 2014;30(3):335-42.
29. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, et al. SWISS-MODEL: homology modelling of protein structures and complexes. 2018;46(W1):W296-W303.
30. Colovos C, Yeates TOJPs. Verification of protein structures: patterns of nonbonded atomic interactions. 1993;2(9):1511-9.
31. Zhang Y, Skolnick JJNar. TM-align: a protein structure alignment algorithm based on the TM-score. 2005;33(7):2302-9.
32. Szklarczyk D, Kirsch R, Koutrouli M, Nastou K, Mehryary F, Hachilif R, et al. The STRING database in 2023: protein–protein association networks and functional enrichment analyses for any sequenced genome of interest. 2023;51(D1):D638-D46.
33. Cotto KC, Wagner AH, Feng Y-Y, Kiwala S, Coffman AC, Spies G, et al. DGIdb 3.0: a redesign and expansion of the drug–gene interaction database. 2018;46(D1):D1068-D73.
34. Kozakov D, Hall DR, Xia B, Porter KA, Padhorny D, Yueh C, et al. The ClusPro web server for protein–protein docking. 2017;12(2):255-78.
35. Bitencourt-Ferreira G, de Azevedo WFJDsfdd. Docking with swissdock. 2019:189-202.