Journal of Population Therapeutics & Clinical Pharmacology

IDENTIFICATION OF SMALL LIGAND MOLECULES AS POTENT INHIBITORS OF THE MYL9 PROTEIN: AS FUTURE CANCER LEADS.

Priyadarshini Gangidi[1], Mounika Badineni[1], Madhavi Latha Bingi[1], Vani Kondaparthi[1], Hareesh Reddy Badepally[1], Kiran Kumar Mustyala[3], Vasavi Malkhed[1,2*]

[1]Molecular Modelling Research Laboratory, Department of Chemistry, Osmania University, Hyderabad, Telangana, India –500007;
[2*]Department of Chemistry, University College of Science, Saifabad, Osmania University, Hyderabad, Telangana, India –500004;
[3]Department of Chemistry, Nizam College, Osmania University, Hyderabad, Telangana, India – 500001

*Corresponding Author: Vasavi Malkhed
*Department of Chemistry, University College of Science, Saifabad, Osmania University, Hyderabad, Telangana, India –500004;

Abstract

A collection of illnesses collectively known as Cancer arises when cells proliferate uncontrollably and spread to other parts of the body. Cancer can impact any tissue or organ. Myosin regulatory light chain 9 (MYL9) or Myosin light polypeptide 9, is the Myosin regulatory subunit that plays a vital role in regulating both smooth muscle and non-muscle cell contractile activity via phosphorylation, implicated in cytokinesis, receptor capping, and cell locomotion. The current study considers the MYL9 protein a prognostic marker and therapeutic target for various tumours. The current investigation aims to identify ligands that may act as inhibitors of the MYL9 protein for the treatment of multiple types of Cancer. The MYL9 protein's homology model is produced using Google ColabFold's AlphaFold. The protein's three-dimensional structure has been verified. The top ten ligands are suggested as effective inhibitors of MYL9 for Cancer treatment based on virtual screening experiments. It has been identified that the amino acid residues ILE41, ILE49, LEU54, MET73, GLU76, MET89, and LYS93 preferentially dock with ligands, yielding a good Glide Score. The proposed ligands are validated through MM-GBSA, SASA, and ADMET properties. The pharmacophore groups that strongly inhibit the MYL9 protein are proposed.

1. INTRODUCTION

The MYL9 protein levels are elevated in various tumors and contribute to tumor invasion in the breast, liver, glioblastoma, pancreatic ductal adenocarcinoma, ovarian epithelial tumors, and esophageal squamous cell carcinoma[1-5]. The MYL9 protein belongs to the EF-hand calcium-binding protein family, specifically the myosin regulatory light chain (RLC) subfamily, with a focus on the functions of smooth muscle and non-muscle cells.

Myosins are a class of motor proteins that use ATP for chemical energy to produce force and motion along actin filaments and are involved in several biological functions, such as cell motility, intracellular transport, and muscle contraction, in the cells of different tissues[6]. The regulatory myosin light chain (MLC), called MLC-2 or MYL9, is responsible for smooth muscle and nonmuscle

cell contractile activity[7]. The MYL9 protein is a protein-coding gene that plays a role in various cellular functions, including vascular function, cancer progression, and smooth muscle contraction [2, 8]. It functions as the regulatory light chain of the myosin protein. According to studies, the MYL9 protein can alter the immunological microenvironment and stimulate the migration, invasion, and angiogenesis of cancer cells [9-11]. The current research considers the MYL9 protein as a prognostic target in cancer treatment.

The MYL9 protein converts chemical energy from ATP into mechanical energy, facilitating the contraction of smooth muscles within the cell. The enzyme Myosin Light Chain Kinase (MLCK) phosphorylates the MYL9 protein, thereby activating it to bind to actin filaments, which governs smooth muscle contraction, cell adhesion, and cell migration. Rho-associated Kinase (ROCK) also promotes phosphorylation of the MYL9 protein. Calcium binding to calmodulin activates MLCK and ROCK, and phosphorylates the MYL9 protein[12, 13]. This promotes cytoskeletal remodeling, thereby expanding the ability of cells to proliferate, differentiate, adhere, and migrate (Figure 1). The present research considers the MYL9 protein as a potential prognostic target in the treatment of various cancers. Inhibiting its function at the calcium-binding site with a potential inhibitor can prevent its activation, thereby preventing cancer progression and metastasis.
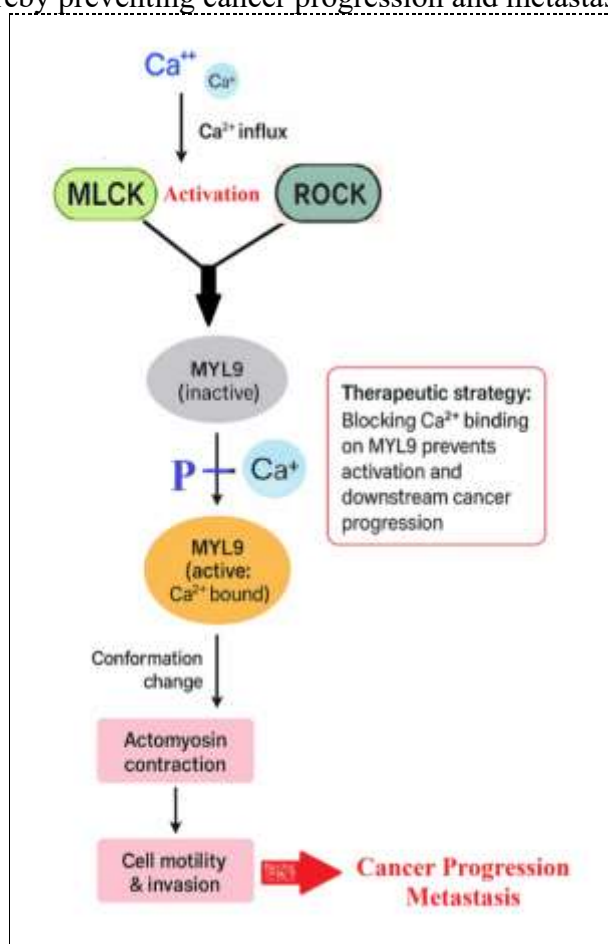

Figure 1. The biochemical pathway of the MYL9 protein in various cancers. Calcium influx activates the MLCK and ROCK systems, which in turn phosphorylate MYL9. The calcium-bound and phosphorylated MYL9 becomes active, contributing to actomyosin contraction, cell motility, and cancer progression.

2. METHODOLOGY and SOFTWARE

Structure-based drug design (SBDD) is a computational technique that makes use of the three-dimensional structures of biological targets, mainly proteins, to guide the discovery and optimization of new therapeutics [14]. By examining the binding sites and structural features of the target, SBDD enables the rational design of ligands that can specifically interact with and modulate protein function.

Techniques such as molecular docking, molecular dynamics simulations, and virtual screening are often employed to predict binding affinities and refine lead compounds. This approach accelerates drug discovery, reduces experimental costs, and has been widely applied in the development of treatments for cancer, infectious diseases, and metabolic disorders[15, 16].

2.1. The MYL9 protein 3D model through the ColabFold

AlphaFold from Google ColabFold is used to create a three-dimensional (3D) model of the MYL9 protein[17]. UniProtKB/Swiss-Prot is the primary source for comprehensive protein sequence annotations. The FASTA sequence of the MYL9 protein, derived from the Universal Protein Resource, is submitted to the ColabFold, which generates the protein's three-dimensional structure[18]. Five 3D models of the protein are generated from AlphaFold, but the 3D model with the highest per-residue local confidence (pLDDT) score, which ranges from 0 to 100, is considered the most stable form of the protein. A higher pLDDT score, closer to 100, indicates greater confidence in the model. Furthermore, the degree of resemblance between the theoretical and real structures is also evaluated using the Template Modelling (pTM) score. The generated 3D model is said to be quite similar to the protein's real structure if the pTM score is higher than 0.5, which goes from 0 to 1.

ColabFold

Google ColabFold's AlphaFold artificial intelligence system applies deep learning to generate highly precise predictions of protein 3D structures. AlphaFold represents the first computational approach capable of reliably forecasting protein conformations at atomic-level accuracy, even in the absence of comparable templates. The performance of AlphaFold has been confirmed through the 14th Critical Assessment of Protein Structure Prediction (CASP14) [19].

In AlphaFold, pairwise residue representation is fundamental for 3D protein structure prediction, as it encodes inter-residue distances and interactions required for proper folding. Large DNA sequencing datasets are aligned into multiple sequence alignments (MSAs), where correlated variations among amino acids reveal likely residue–residue contacts. AlphaFold leverages these MSAs to detect evolutionary constraints that encode structural information. By learning from these patterns, the ColabFold predicts accurate 3D protein structures [20]. A contact is typically defined as the presence of β-carbon atoms from two residues within a distance of 8 Å. Neural networks estimate contact probabilities from MSA-derived features and incorporate these into structural modeling by adjusting statistical potentials, thereby steering folding toward conformations that satisfy predicted interactions. Additionally, pairwise distances and backbone torsion angles are expected to refine the final structure. To improve accuracy, evolutionary distance-based potentials derived from neural network predictions, independent of covariation features, are applied to rank models.

2.2. Verification of the MYL9 protein 3D model

Verifying the 3D structure of a computationally generated protein model is crucial for target-driven drug discovery, as researchers must select the most suitable tertiary structure models from a large number of predicted models before employing them in biomedical studies. It helps comprehend the protein's relationships, function, and involvement in drug design. Verifying the protein model's correctness helps avoid mistakes by ensuring it matches the actual structure [21]. Protein Structure Analysis (ProSA), Predicted Aligned Error (PAE) heatmap, and Ramachandran Plot (RP) are used to assess the generated 3D protein model. ProSA is an interactive web service that researchers use to check errors in protein 3D structures through quality scores and energy plots [22]. ProSA identifies the problematic areas in the 3D model generated, ensuring the researchers have reliable models. A significant outcome is the Z-score, which shows the model's overall quality and is presented alongside all of the PDB's known protein structures [23]. A negative Z-score is regarded as favorable, as it suggests fewer errors in the protein model. In structural analysis, the ProSA local quality graph illustrates knowledge-based energy profiles relative to amino acid sequence positions, enabling the identification of problematic or misfolded regions within a protein [24].

The PROCHECK package evaluates torsion angles (φ and ψ) through the Ramachandran Plot (RP). The plot is a two-dimensional representation where the φ angle is plotted along the x-axis and the ψ angle along the y-axis [25]. The φ and ψ values for each amino acid residue in a structure are indicated as points, showing whether the torsion angles are stereochemically feasible [26]. Additionally, the Predicted Aligned Error (PAE) heatmap generated by the ColabFold displays the predicted residue-level error relative to reference structures [27].

2.3. Identification of Binding Sites in the MYL9 protein
A binding site or active site in a protein is defined as the specific amino acid residues where small molecules (ligands) form strong interactions, such as hydrogen bonds. The identification is crucial in drug design because it helps hinder protein activity by blocking active amino acid residues or inducing conformational changes around the active site. The characteristics of the site define the features and binding potential of the ligands identified.

The CASTp (Computed Atlas of Surface Topography of Proteins) server [28] and the SiteMap suite of Schrödinger tools estimates a protein's binding site. CASTp is a web-based tool designed to locate, define, and quantify concave surface areas on three-dimensional protein structures, such as voids and pockets [29]. It analyzes geometric and topological aspects using the alpha shape approach from computational geometry, which estimates the location and volume of regions using models such as the solvent-accessible surface (Richards' surface) and the molecular surface (Connolly's surface) [30]. For pre-computed results, a standard probe size of 1.4 Å is applied to calculate solvent-accessible surface area, approximating the dimensions of water molecules [31]. Within structure-based drug design, Schrödinger's SiteMap program is commonly employed to locate, characterise, and evaluate the druggability of possible binding sites on protein surfaces [32, 33]. The tool embeds the protein in a 3D Cartesian grid labelled for the atom's Van der Waals radius.

2.4. Virtual Screening of the MYL9 protein
Virtual screening is a computational approach used in drug discovery to identify ligands or small molecules with the potential to interact with the target protein and act as effective inhibitors [34]. The process prepares the ligands and target protein for docking and scoring analysis[35]. Ligands are filtered based on binding scores. Biochemical assays and crystallography experimentally validate the identified ligands.

Ligand Preparation
In the ligand preparation process for the MYL9 protein molecular docking, the ligand's configuration is adjusted to be suitable for precise docking simulations [36]. This process includes generating four conformers for ligands at pH 7.0 ± 2.0, adding missing hydrogens, assigning partial charges, and defining rotatable bonds to ensure proper ligand interaction at the receptor's active site, where the lowest-energy conformer is selected for further study [37, 38]. The preparation of ligand structures for virtual screening was performed using the LigPrep module of the Schrödinger package, employing the OPLS force field [39].

The MYL9 protein Preparation
The Protein Preparation module in the Schrödinger Suite optimises protein structures for energy minimisation, ensuring high-quality structures for docking and simulations in drug design by adjusting atomic coordinates and reducing potential energy[40]. During this process, missing hydrogens are added, side chains are corrected, and hydrogen bond networks are optimised to ensure optimal functionality. The Schrödinger Suite utilises the OPLS force field for energy minimisation. It determines a protein's potential energy, which is essential for replicating experimental findings [41, 42].

Receptor Grid Generation in the MYL9 protein binding site

In molecular docking, receptor grid development entails mapping a target protein's interaction potential with ligands by constructing a three-dimensional grid around the protein's binding site. The grid specifies spatial coordinates to evaluate ligand atoms for energy interactions, such as van der Waals and electrostatic forces [43]. Through the pre-computation of receptor characteristics, this procedure reduces computational complexity and facilitates the adequate sampling of ligand poses during docking. The glide module of the Schrödinger suite is used for grid generation and effective ligand-receptor docking, and it is helpful in virtual screening, binding pose modification, and pose optimization [44]. Glide utilizes the OPLS-AA force field and hierarchical filters to provide accurate energy estimates. The Schrödinger suite generates the Grid using the Grid Gen module, with dimensions of 80 Å × 80 Å × 80 Å for the 3D structure generated. This grid encloses the cavity of the active site, guaranteeing every ligand's potential for binding[45].

Molecular Docking of the MYL9 protein

Molecular docking is a computational method applied in target-based drug discovery or SBDD to examine atomic-level interactions between ligands and the receptor protein (MYL9) [46]. It serves as a practical strategy in drug design and development, assisting in the identification and optimization of potential therapeutic compounds. The receptor–ligand docking was performed using the GLIDE (Grid-based Ligand Docking) software, where the protein Grid defined the active binding pocket.

Three phases of filtration and inquiry are used to analyse the ligand structures; each stage yields an output of 10%. The High Throughput Virtual Screening (HTVS) protocol was then applied, which systematically screens the ligand library to eliminate unstable molecules and retain low-energy, flexible ligand conformations. Monte Carlo simulations [47] examine the structural and orientational isomers at the Standard Precision (SP) stage, when ligands with a high binding affinity for the receptor group bind together. SP outperforms random selection in 97% of situations, reproducing crystal complex geometries with <2.5 Å RMSD in 85% of Astex cases and achieving an average AUC of 0.80 across 39 DUD targets. At the Extra Precision (XP) step, ligand structures are screened according to energy and scoring function criteria. The XP (Extra Precision) method was developed to improve screening accuracy and better reproduce experimental binding data. It achieves docking pose accuracies with RMSDs around 2 Å for well-docked ligands and root-mean-square errors of ~2 kcal/mol in predicted binding affinities [48]. Conformers showing the best fit, optimal energy values, and highest Glide scores, as obtained from the extra precision mode, are prioritised for further investigation. The Glide Score is a scoring function designed to estimate the strength and stability of ligand-receptor interactions during molecular docking studies. It evaluates how well a ligand fits into the binding pocket of a target protein by considering both favourable interactions, such as shape complementarity, electrostatic attraction, hydrophobic contacts, and hydrogen bonding, as well as unfavourable factors, including excessive flexibility or poorly oriented polar groups. The Glide score reflects the strength of interaction between the ligand and receptor. The screened ligands are processed for validation [49].

2.5. Validation of Ligands

In experimental or computational investigations, ligand validation ensures that a ligand, a molecule that binds to a specific site on a macromolecule (such as a protein or receptor), is accurately represented in terms of its structure, binding capabilities, and functional significance. This guarantees the ligand's reliability for applications such as drug development, structural biology, or biochemical research [50].

MM-GBSA

The binding strengths between protein and ligand estimate the strength of the docking. The binding free energy of protein-ligand complexes or other biomolecular interactions is estimated computationally in molecular dynamics (MD) simulations using the MM-GBSA (Molecular Mechanics-Generalised Born Surface Area) method [51, 52]. In the present approach, molecular

mechanics (MM) energy terms are combined with solvent-accessible surface area (SA) calculations and continuum solvent models, such as the Generalised Born (GB) approximation, to evaluate the binding free energy [53-55].

ΔG_bind = G_complex - (G_receptor + G_ligand)

Free energy, **G = E_MM + G_solv - TS**

E_MM: Molecular Mechanics energy (internal, van der Waals, electrostatic)

G_solv: Solvation free energy

TS: Entropy contribution

SASA of the docked complexes

The Solvent-Accessible Surface Area (SASA) of docked complexes represents the portion of a protein or protein–ligand complex exposed to solvent molecules, typically water [56]. SASA provides insights into hydrophobic compactness, conformational stability, and folding of proteins, along with their interactions with ligands or other biomolecules [57]. In the context of docked complexes, SASA is widely used to evaluate conformational changes, binding interface stability, and the effect of mutations on protein-ligand and protein-protein interactions [58]. Schrödinger software, through its Maestro interface, offers SASA analysis. The Shrake-Rupley algorithm approximates SASA by placing a probe sphere, typically 1.4 Å in radius, similar to a water molecule, around the molecule[59]. It calculates the Surface Area accessible to the probe without overlapping with other atoms.

ADMET studies of the docked ligands

In drug development, ADMET (Absorption, Distribution, Metabolism, Excretion, and Toxicology) studies of ligands are crucial for evaluating how the compounds behave in living organisms and how that affects their potential as therapeutic agents [60]. The Schrödinger QikProp module calculates the ADME properties of the ligands docked with the MYL9 protein[61]. QikProp evaluates over 20 physicochemical descriptors, including HERG potassium channel inhibition (log IC50), Caco-2 and MDCK evaluations of membrane permeability, blood–brain barrier penetration (log BB), water solubility (log S), octanol–water partition coefficient (log P), and human serum albumin binding (log Khsa) [62, 63]. The forecasts reduce the clinical trial failures, facilitating the early identification of drugs with favourable pharmacokinetic profiles. A drug's detrimental effects on the body as a result of excessive exposure, improper dosage, or unfavourable reactions are referred to as drug toxicity. The toxicity of the drug depends on the ability to metabolise and excrete the drug [64, 65].

3. RESULTS AND DISCUSSION

3.1. AlphaFold 3D model of the MYL9 protein generation

The 3D structure of the MYL9 protein is generated **(Figure 2)** by employing ColabFold. The 172-amino-acid FAASTA sequence, which has the UniProt ID P24844, is obtained from UniProtKB and uploaded to ColabFold. For additional analysis, the 3D model with pLDDT = 83.10 and pTM = 0.512 is used. A pLDDT score of 83.10 on a scale of 0-100 and a pTM score of 0.512 indicate that the generated model is reliable.
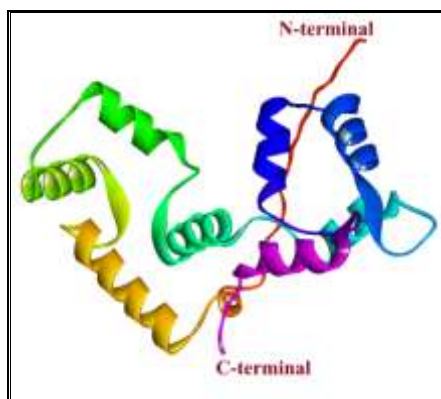

Figure 2. ColabFold-generated 3D structure of the MYL9 protein, yielding a pLDDT score of 83.10 and a pTM score of 0.512. It is visualised in Discovery Studio.

3.2. Evaluation of the ColabFold-generated 3D model of the MYL9 protein

The ProSA server was employed to validate the constructed 3D model of the MYL9 protein, which produced a Z-score of –6.03 (Figure 3). This result suggests that the predicted model of MYL9 has fewer structural errors compared to many experimentally resolved structures available in the PDB. The local model quality profile (Figure 4) further shows that all amino acid residues fall within the negative energy region (below zero), indicating that the MYL9 structure is energetically stable and structurally reliable, thereby enhancing the overall quality of the model.
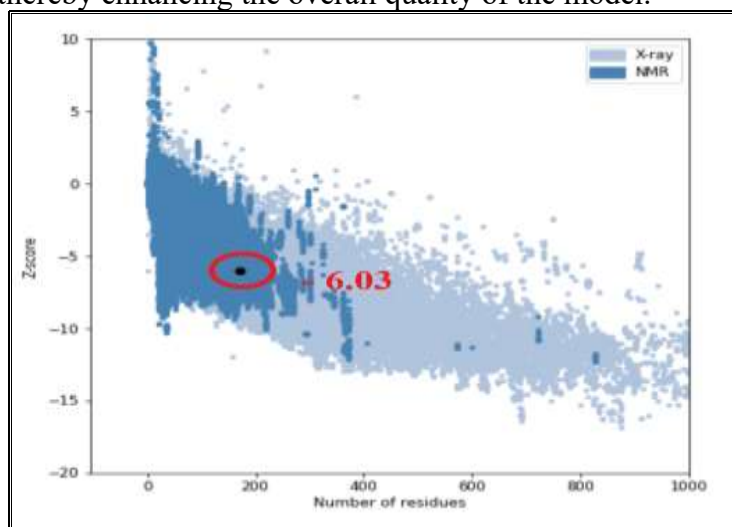

Figure 3. ProSA validation of the MYL9 protein model. The predicted 3D structure of MYL9 shows a Z-score of –6.03, which is consistent with experimental protein structures in the PDB, indicating fewer structural errors. In the ProSA plot, the light blue region corresponds to X-ray data and the dark blue region to NMR data of experimentally resolved proteins in the PDB. The location of the MYL9 protein is highlighted by a black dot inside a red circle.
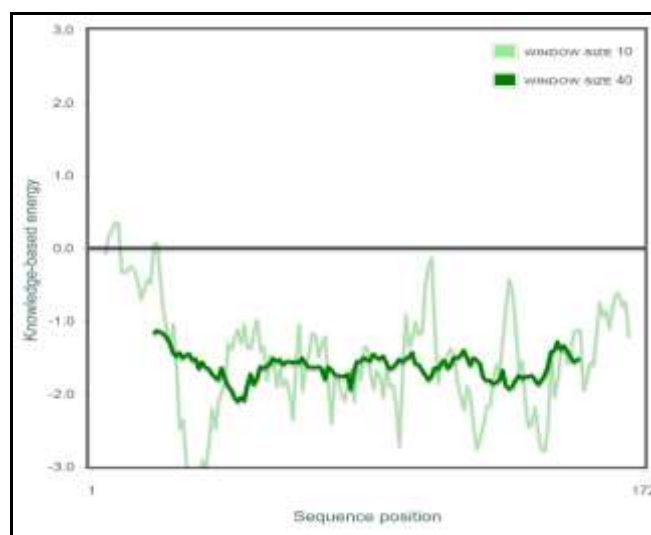

Figure 4. The amino acid residues in the negative, less energetic area are displayed on the ProSA-assessed local model quality graph for the MYL9 protein. The negative energy indicates that the amino acid configuration is acceptable.

The 3D model of MYL9 has been validated using the Ramachandran Plot (RP Plot), as shown in Figure 5, generated by the PROCHECK server. The RP Plot illustrates the overall stereochemical feasibility of the constructed MYL9 protein model. It features four distinct regions: the energetically most favoured zone, the extra permitted region, the liberally allowed region, and the forbidden region (Figure 5A). According to the plot data, 87.0% of the residues (134 amino acids) in the MYL9 protein fall within the most preferred region, while 13% (20 amino acids) reside in the extra allowed areas.

Notably, no residues are located in the liberally permitted or disallowed regions. Considering that 87.0% of the amino acid residues are positioned in the most preferred region, it can be concluded that the model demonstrates good stereochemical feasibility (Figure 5B).
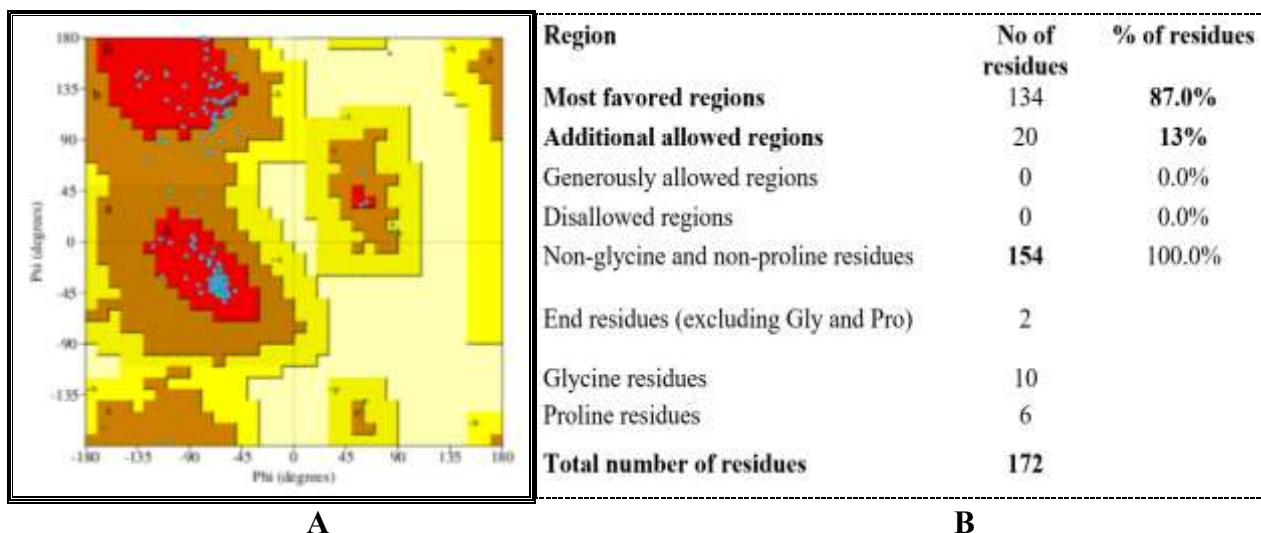

Region	No of residues	% of residues
Most favored regions	134	**87.0%**
Additional allowed regions	20	**13%**
Generously allowed regions	0	0.0%
Disallowed regions	0	0.0%
Non-glycine and non-proline residues	154	100.0%
End residues (excluding Gly and Pro)	2	
Glycine residues	10	
Proline residues	6	
Total number of residues	172	

           A   **B**

Figure 5. The MYL9 protein Ramachandran plot. **(A)**Four areas make up the account. The red region shows the most preferred area, the additionally permitted area by the brown region, the liberally permitted area by the yellow region, and the prohibited area by the light yellow field. **(B)**The proposed model appears stereochemically acceptable since 134 amino acids (87.0%) of the 172 amino acids in the MYL9 protein are located in the most preferred area.

The PAE heat map generated from the ColabFold (Figure 6) shows clear patterns of structural confidence across the protein. The dark green region indicates low error and high confidence, and the calcium-binding region 30-165 falls under this region.
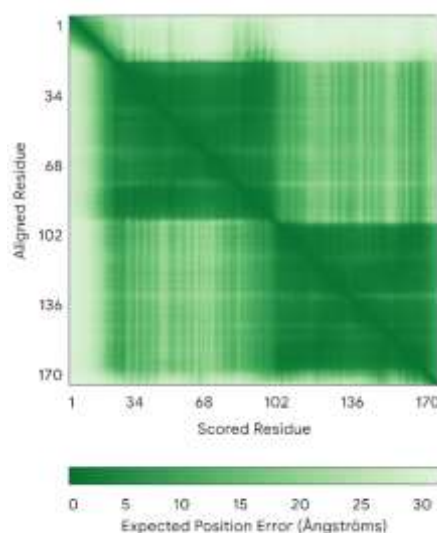

Figure 6. PAE plot of the MYL9 protein. The dark green region indicates the low PAE value with high confidence, while the pale green region indicates low confidence with a high PAE value.

### 3.3. The secondary structural characteristics of the MYL9 protein
Tables 1 and 2 present the composition of the MYL9 secondary structure, which includes α-helices and β-sheets. The protein contains 9 α-helices, 2 β-sheets, 7 β-strands, and 2 β-hairpins. Pi-cation interactions, along with intramolecular hydrogen bonds, play a key role in maintaining the structural integrity of MYL9. The secondary structure information of MYL9, obtained from the PDB-Sum server, is illustrated in Figure 7.

Table 1.α-helices of the MYL9 protein.

S.no	Starting amino acid Residues	Ending amino acid Residues	No. of residues	Length in Aº	Sequence
1	VAL22	MET25	4	7.40	VFAM
2	GLN28	ILE41	14	21.37	QSQIQEFKEAFNMI
3	LYS51	SER60	10	15.68	KEDLHDMLAS
4	ASP67	SER75	9	14.05	DEYLEGMMS
5	PHE83	LEU94	12	17.91	FTMFLTMFGEKL
6	GLU100	ALA108	9	13.73	EDVIRNAFA
7	GLU120	THR129	10	15.02	EDHLRELLTT
8	ASP136	TYR143	8	12.56	DEEVDEMY
9	TYR156	HIS165	10	14.23	YVEFTRILKH

Table 2.β-Strands of β-Sheets of the MYL9 protein.

S.no	Starting amino acid Residue	Ending amino acid Residue	Beta Sheet	No.of residues	Sequence
1.	PHE48	ILE49	A	2	FI
2.	ILE81	ASN82	A	2	IN
3.	PHE117	HIS119	B	3	FIH
4.	ASN153	ASN155	B	3	NFN
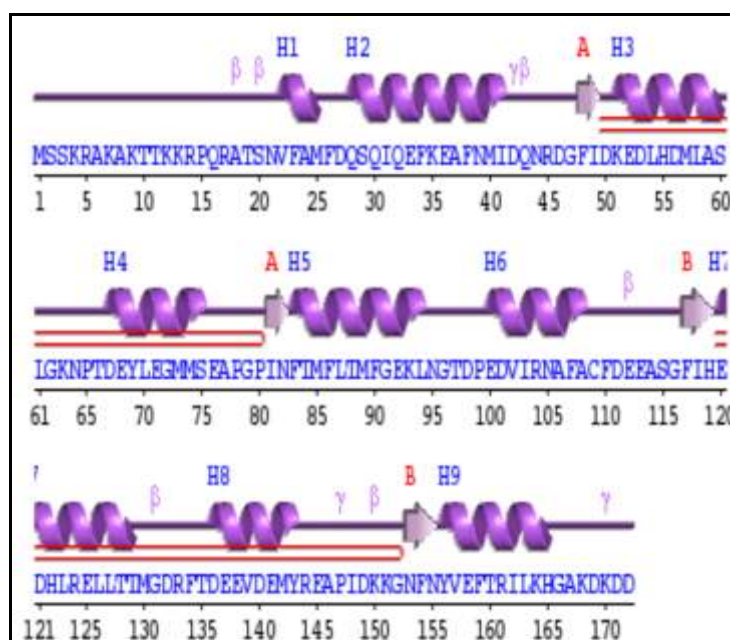
Figure 7. The schematic diagram of the secondary structure of the MYL9 protein from the PDB-Sum server.

3.4. Binding site prediction of the MYL9 protein

The binding pocket of the MYL9 protein was predicted using the SiteMap tool in the Schrödinger suite, in conjunction with the CASTp server. The calculated pocket volumes were 167.041 Å³ and 687.78 Å³ for Site 1 (CASTp) and Site 2 (SiteMap), respectively. These regions, which contain amino acid residues essential for ligand interaction, are summarized in Table 3. Conserved residues spanning positions 30 to 165, as identified through the BLASTp algorithm (Figure 10), overlapped with residues mapped to the predicted binding pockets by SiteMap (Figure 8) and CASTp (Figure 9). Based on these predicted binding regions, Schrödinger's Grid Generation tool was employed to prepare the MYL9 receptor for Glide docking.
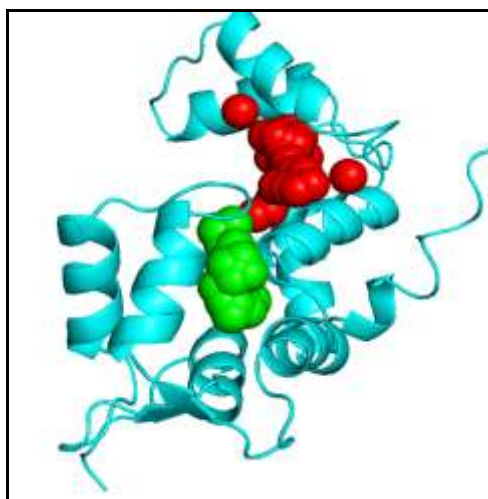
Figure 8. Putative binding pockets 1 and 2, respectively, in red and green, were generated from the SiteMap tool and viewed from PyMOL.
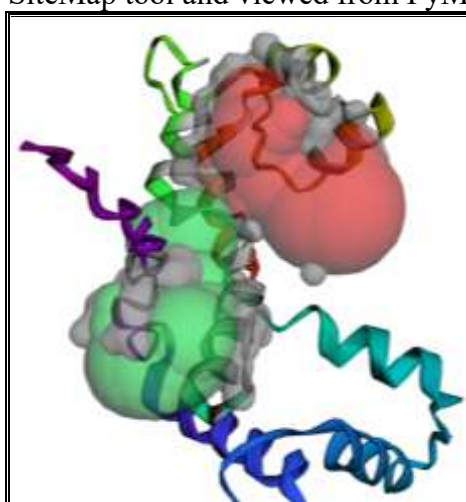

Figure 9. The putative binding pockets 1 and 2, respectively, are shown in red and green, as viewed from the CASTp server.
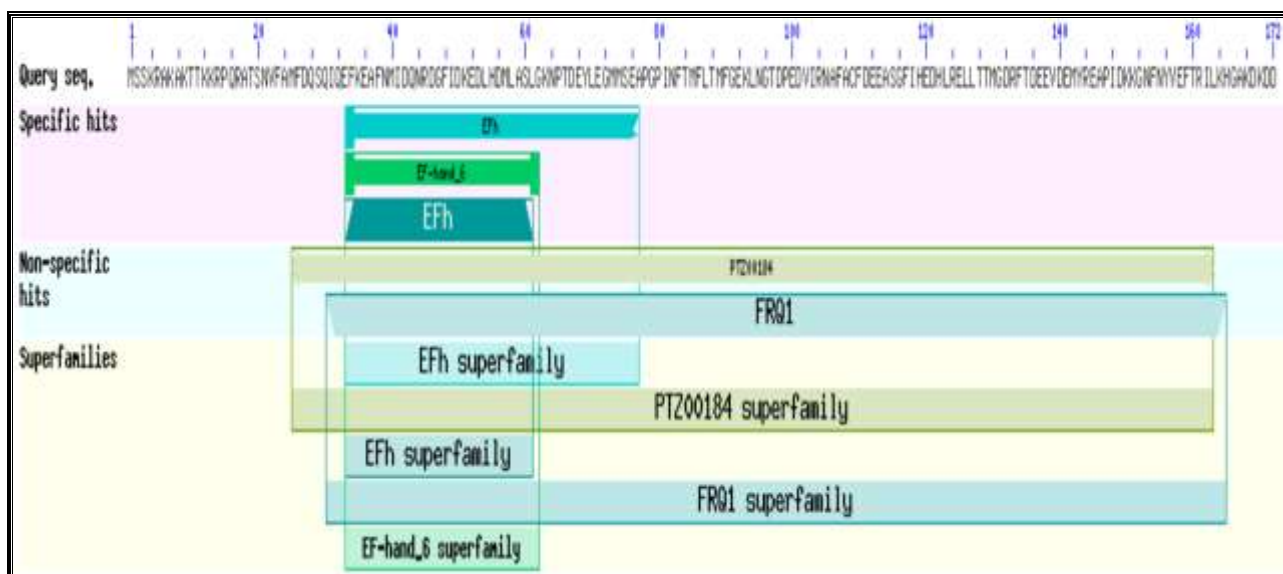

Figure 10. The MYL9 protein conserved domains obtained via BLASTp. The calcium-binding site and active amino acid residues 30–165 are binding sites for new lead compounds.
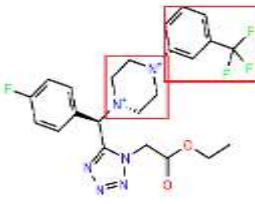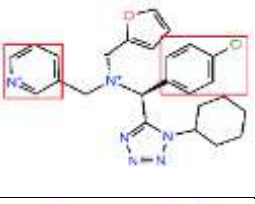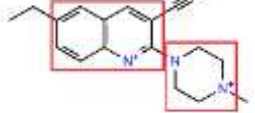
Table 3. Volume and Amino acid residues present in the binding sites of the MYL9 protein from the CASTp tool and the Site Map tool.
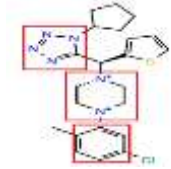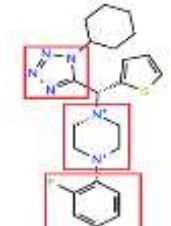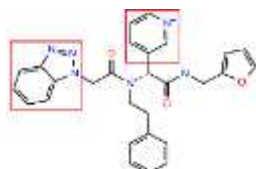
S.No	Server	Cavity	Volume in Å³	Residues
1.	CASTp	1	687.78	93LYS, 96GLY, 97THR, 89ASP, 102VAL, 106ALA, 107PHE, 110PHE, 118ILE, 123LEU, 126LEU, 127LEU, 130MET, 131GLY, 132ASP, 133ARG, 134PHE, 138GLU, 142MET, 154PHE, 159PHE, 163LEU
		2	105.6	144ARG, 145GLU, 146ALA, 147PRO, 158GLU, 161ARG, 162ILE, 165HIS, 166GLY, 176ALA, 168LYS, 170LYS
2.	Site Map	2	167.041	ILE41, ILE49, LEU54, MET73, GLU76, ALA77, PRO78, ILE81, PHE86, MET89, PHE90, GLU92, LYS93
		1	120.393	PHE107, PHE110, ILE118, LEU123, LEU126, LEU127, MET130, GLY131, ASP132

3.5 Virtual Screening Analysis of the MYL9 protein

Schrödinger's LigPrep module prepares the ligand dataset for docking with the MYL9 protein in the Virtual Screening process. Then, the protein is optimized using the Protein Preparation Wizard. A grid is generated to encompass the binding site of the MYL9 protein, ensuring accurate docking. The ligand dataset is submitted to the Virtual Screening window, which performs High-Throughput Virtual Screening, filtering out low-energy, flexible ligands for Standard Precision Screening. Standard Precision Docking then selected stable structural orientations of isomers. Next, the chosen structures undergo Extra Precision (XP) screening, where ligands are further filtered based on energy and scoring criteria. The conformers with the most negative glide scores and energies are the ones that dock most closely to the MYL9 protein. Fifty-nine ligands are selected after XP screening, with the ten best-fit ligands presented in Table 4, where the interacting pharmacophore of ligands is highlighted in a red box. Preferably, the MYL9 protein docked with the amino acids ILE41, ILE49, LEU54, MET73, GLU76, MET89, and LYS93 to the ligands. The important pharmacophore groups with which the ligands docked with the MYL9 protein are Trifluoromethylphenyl, Indole, Piperazine, Chlorophenyl, Pyridine, quinolone, fluoro-benzyl, methoxy-benzyl, tetrazole, benzotriazole, quinolin-2-one, and dimethylamino-phenyl. The 3D and 2D images of the target-ligand complexes are shown in Figure 11. These amino acids fall in the EF-hand motif, which functions as a calcium-binding and signal transduction.

Table 4. The selected ligands from L1 to L10 and their interactions with the target protein, MYL9

Ligand .no	Structure of the molecule	Glide Energy (Kcal/Mol)	XP Glide Score	Intermolecular Interaction	Distance (Å)
L1		-49.74	-8.96	L1: H25 - A: GLU76:OE2 L1: H11 - A: MET73: O L1 - A: LYS93 L1:C2 - A: ILE41 L1 - A: LEU54 L1:C22 - A:PRO78 L1:H20 - A: GLU92:OE2	1.77 2.44 4.03 4.69 5.29 3.97 2.23
L2		-47.31	-8.88	L2: H28 - A: GLU76: OE2 L2: H29 - A: MET89: O L2: CL1 - A: LYS93 L2 - A: MET73 A: PHE90 -: L2:CL1	2.15 1.73 4.47 5.09 4.19
L3		-33.30	-8.70	L3: N3 - A: GLU76: OE2 L3: N4 - A: GLU92: OE2 L3: H2 - A: MET89: O L3: C16 - A: ILE49 L3: C16 - A: ILE41 L3: C16 - A: LEU54	3.84 4.37 2.33 4.35 4.36 4.56

|					
L4		-52.78	-8.42	L4: H1 - A: GLU76: OE2 A: LYS93:HZ1 -: L4:O2 A: LYS93:HE2 -: L4: O1 L4: H11 - A: MET73: O L4 - A: MET89 L4:H11 - A: MET73: O L4 - A: LEU54 L4 - A: ILE81	1.79 1.80 2.57 2.24 5.22 2.24 5.49 5.20
L5		-47.32	-8.37	L5: H30 - A: GLU76: OE2 A: LYS93: HZ1 -: L5:N4 L5: H8 - A: MET73: O A: MET89 -: L5 L5:CL1 – A: ILE41 L5:C18 – A :LEU54	1.74 2.16 2.57 5.39 5.49 4.86
L6		-47.94	-8.27	L6: H26 - A: GLU76:OE2 A: LYS93: HZ1 -: L6: N4 L6: H7 - A: MET73: O A: MET89 -: L6 L6 - A: LEU54	1.65 2.42 2.49 5.11 5.34
L7		-50.76	-8.20	L7: H6 - A: GLU76: OE2 A: LYS93: HE2 -: L7: O3 L7: H14 - A: MET89: O L7 - A: MET73 L7 - A: ILE41 L7 - A: ILE49 L7 - A: LEU54	1.69 2.50 2.86 5.05 5.44 5.30 5.16
L8		-43.04	-8.14	L8: H35 - A: GLU76: OE2 L8: C17 - A: LYS93 L8: C19 - A: MET89 A: PHE90 -: L8 L8:C19 – A: ILE81 L8:C17 – A: LEU94 L8:C21 – A: ILE41 L8:C21 – A: ILE49 L8:C21 - A: LEU54	1.67 4.04 4.88 3.86 4.64 4.17 4.53 4.23 5.20
L9		-49.49	-8.06	L9: H37 - A: GLU76: OE2 A: PHE90 -: L9 L9 - A: MET73 L9 - A: MET89 L9 - A: LYS93 L9:C18 – A: LYS93 L9:C18 – A: LEU94 A: ILE49 -: L9 A: LEU54 -: 2L9	1.69 3.93 5.38 4.72 5.49 3.97 3.89 4.99 5.34
L10		-49.20	-8.00	L10:N4 - A: GLU76: OE2A: A: LYS93: HZ1 -: L10: O1 L10:H2 - A: MET73: O: L10 L10 : H8 - A:MET89:O L10 - A: ILE49 L10 - A: LEU54 L10 - A: ILE81	3.92 2.38 1.76 2.61 5.35 5.07 4.69
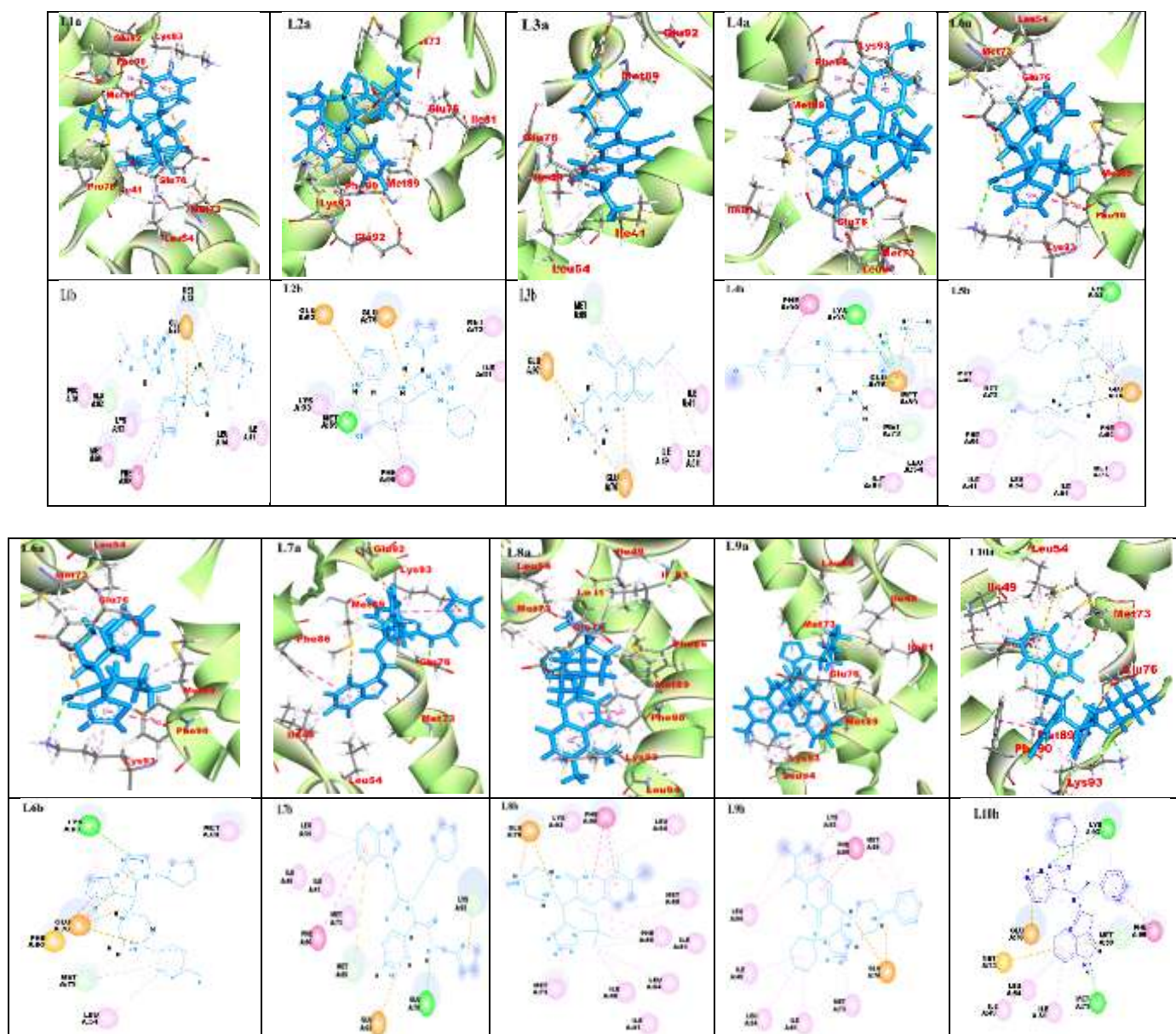
Figure 11. 3D structures of the MYL9 protein docked with ligands L1a to L10a. The docked ligands are depicted in a stick model near the protein, which is represented as a flat ribbon. The amino acids interacting with the ligands are displayed. 2D structures of the ligands and their interactions with the protein are shown in L1b to L10b. The interacting amino acid residues are represented as balls. N-π, π-alkyl, and H-bond interactions are indicated in orange, pink, and green, respectively. The 3D and 2D structures of the docked complexes are visualized using Discovery Studio.

3.6. MM-GBSA calculations of the Ligands:
The binding free energies of protein-ligand complexes were calculated from the Prime MM-GBSA module of the Schrödinger docking suite. Table 5 provides the computed outcomes obtained through the MM-GBSA tool available in the Schrödinger suite. The approach evaluates different components of binding energy, such as hydrogen bonding, covalent bonding, Van der Waals forces, and solvation contributions. Negative energy values indicate the thermodynamic stability of the docked ligands with the MYL9 protein. The priority given is based on the Glide Score.

Table 5. MM-GBSA analysis of the MYL9-ligand complexes.

Ligand No	XP Glide score	MM-GBSA dG bind (k.cal/mol)	MM-GBSA dG Bind Coulomb (K.cal/mol)	MM-GBSA dG Bind Covalent (k. cal/mol)	MM-GBSA dG Bind vdW (k.cal/mol)	MM-GBSA dG Bind Solv GB (k. cal/mol)	MM-GBSA dG Bind H-bond
L1	-8.96	-45.43	-151.36	8.69	-39.46	159.96	-1.28
L2	-8.88	-46.05	-211.09	3.86	-37.28	217.56	-1.06
L3	-8.70	-44.24	-192.24	0.68	-28.78	194.20	-0.02
L4	-8.42	-43.42	-157.86	5.49	-33.63	166.06	-2.01
L5	-8.37	-51.59	-155.58	8.20	-42.88	163.99	-2.01
L6	-8.27	-33.74	-159.08	10.53	-22.28	163.68	-0.69
L7	-8.20	-31.40	-150.81	16.10	-42.32	162.39	-0.84
L8	-8.14	-36.09	-141.69	10.47	-27.36	146.71	-0.76
L9	-8.06	-37.39	-150.17	10.05	-32.45	158.72	-0.39
L10	-8.00	-39.50	-143.73	5.86	-33.97	153.45	-1.06

Among the ten ligands (L1–L10), docking and MMGBSA analysis identified L5 and L10 as the most promising candidates. L5 showed a highly favorable binding free energy (–51.59 kcal/mol) with strong van der Waals and hydrogen bonding contributions, while L10 demonstrated the most favorable overall binding energy (–53.58 kcal/mol), driven by strong electrostatic interactions but has high solvation energy. L1 achieved the best docking score (–8.96 kcal/mol) with stable van der Waals interactions, and L2 also exhibited strong electrostatic interactions but faced solvation penalties. Overall, L5 emerged as the most well-rounded ligand, combining strong binding affinity, van der Waals stabilization, and hydrogen bonding, making it a key candidate for further research.

3.7. SASA studies of the MYL9 protein.

An important factor that governs protein folding and structural stability is its solvent exposure, described as the Solvent Accessible Surface Area (SASA). SASA values are predicted from protein sequences through advanced machine learning methods developed using solved structural datasets and computational algorithms. The solvent-accessible area is often employed to estimate variations in free energy when a biomolecule shifts from a polar (aqueous) medium to a non-polar environment, such as a lipid phase. Table 6 provides SASA estimations for ten receptor–ligand complexes. The obtained values, ranging between 300 and 700, suggest that significant energy redistribution occurs during the MYL9 protein's transition, emphasising its solvent accessibility and highlighting the adaptability of the MYL9 protein–ligand system in both polar and nonpolar surroundings.

Table 6. SASA estimations for ten receptor-ligand complexes.

LIGAND No	SASA (Å^2)	FISA (Å^2)	FOSA (Å^2)	PISA (Å^2)	PSA (Å^2)
L1	822.11	135.12	248.66	256.49	93.07
L2	749.90	60.03	249.93	368.33	53.59
L3	578.61	70.22	350.41	157.97	42.90

L4	852.06	83.07	241.11	480.86	104.81
L5	693.01	76.72	320.31	210.05	55.62
L6	654.18	81.61	290.98	251.95	55.08
L7	827.75	112.48	88.53	626.73	109.50
L8	727.34	117.56	510.57	99.20	88.13
L9	828.53	109.17	443.82	275.53	88.97
L10	799.6	78.49	246.60	474.49	71.58

Surface area analysis revealed diverse physicochemical profiles among the ligands. L4, L7, and L9 exhibited the highest solvent-accessible surface areas (SASA), indicating greater solvent exposure, while L3 showed the lowest SASA. Hydrophilic contributions (FISA) were most pronounced in L1, L7, and L8, whereas hydrophobic surface areas (FOSA) were highest for L8 and L9, suggesting strong lipophilic interactions. L7 and L4 displayed the largest π-interaction surface areas (PISA), favoring aromatic stacking, while polar surface area (PSA) was highest for L7 and L4, correlating with better hydrogen bonding and solubility. Collectively, these results suggest that L7 and L4 balance polar interactions and π-stacking, whereas hydrophobic contacts dominate in L8 and L9, and L3 exhibits the lowest polar surface, thereby favoring membrane permeability.

3.8. ADMET properties
The ADMET properties of the MYL9 protein are summarised in Table 7. These characteristics were calculated using the ADMET tool from the Schrödinger Suite. The human oral absorption values for the selected ligands range from 68% to 100%, indicating that these ligands are likely to be well-absorbed and could serve as potential lead molecules. The blood-brain barrier partition coefficients range from -0.10 to 0.94, while the central nervous system (CNS) values fall between -1 and -2. The number of donor hydrogen bonds is between 0 and 2, and the acceptor hydrogen bonds range from 2.5 to 9.75. Additionally, the molecular weights of the ligands vary from 280.37 to 497.64, all within acceptable limits. These properties suggest that the selected ligands are potent inhibitors of the MYL9 protein, as their ADME characteristics fall within a favorable pharmacokinetic range.

Table 7. ADME properties of the MYL9 inhibitors.

Ligand No	XP Glide score	CNS	Mol MW	Donor HB	Accepted HB	Q Plog Po/w	Q Plog BB	QPlog HERG	% Human oral absorption	Rule of Three	Rule of Five
L1	-8.96	1	492.47	0	8	4.41	-0.53	-7.45	90.56	1	0
L2	-8.88	1	462.98	0	7	4.84	0.10	-7.07	100	0	0
L3	-8.70	1	280.37	0	5	2.52	0.16	-5.70	90.54	0	0
L4	-8.42	-1	478.52	2	9.75	3.70	-0.94	-4.94	100	1	0
L5	-8.37	1	443.00	0	6	4.38	0.25	-5.97	100	0	0
L6	-8.27	1	426.55	0	6	3.79	0.10	-5.81	96.05	0	0
L7	-8.20	-2	494.5	1	9.5	3.63	-1.27	-5.33	100	0	0
L8	-8.14	1	423.56	1	9.5	2.01	-0.12	-6.49	68.69	0	0
L9	-8.06	1	497.64	1	8.5	4.46	-0.40	-7.1	95.29	0	0

| L10 | -8.00 | -1 | 480.60 | 2 | 7 | 4.48 | -0.74 | -4.31 | 100 | 1 | 0 |

ADMET profiling indicated that most ligands met Lipinski's drug-likeness criteria, with molecular weights below 500 Da and acceptable hydrogen bond donor/acceptor counts. Ligands L1, L2, L5, and L9 showed favorable LogP values and high predicted oral absorption (>90%), with moderate CNS penetration. L4 and L7 had weaker blood–brain barrier permeability but maintained excellent oral absorption. HERG inhibition data suggested a potential cardiotoxicity risk for L1, L2, L9, and L10, while L4 and L7 appeared safer. Overall, L5 and L2 emerged as balanced candidates, combining strong docking affinity, favorable pharmacokinetics, and good oral absorption. Conversely, L10 demonstrated excellent binding but violated drug-likeness rules, limiting its potential for further optimization.

4. CONCLUSION

The current research suggests that the MYL9 protein is a promising target for treating various types of Cancer. Using ColabFold software, the three-dimensional structure of MYL9 was generated with AlphaFold and validated. Virtual screening revealed that all ligands preferentially bind to ILE41, ILE49, LEU54, MET73, GLU76, MET89, and LYS93 at the selected binding site. Ten small-molecule ligands were identified based on Glide scores and Glide energy. The ligands from L1 to L10 were further validated through MM-GBSA, ADMET, and SASA analyses. These results highlight L5 as the most promising lead, with L2 and L1 as strong secondary candidates for further optimisation and development. Key pharmacophore groups, such as Trifluoromethylphenyl, Indole, Piperazine, Chlorophenyl, Pyridine, quinolone, fluoro-benzyl, methoxy-benzyl, tetrazole, benzotriazole, quinolin-2-one, and dimethylamino-phenyl, are proposed to inhibit MYL9, potentially aiding in the prevention of various Cancer types and supporting the development of new targeted therapies.

REFERENCES
1. Wen, B., et al., *MYL9 promotes squamous cervical cancer migration and invasion by enhancing aerobic glycolysis.* Journal of International Medical Research, 2023. **51**(11): p. 03000605231208582.
2. Lv, M., L. Luo, and X. Chen, *The landscape of prognostic and immunological role of myosin light chain 9 (MYL9) in human tumors.* Immunity, Inflammation and Disease, 2022. **10**(2): p. 241-254.
3. Kruthika, B.S., et al., *Expression pattern and prognostic significance of myosin light chain 9 (MYL9): a novel biomarker in glioblastoma.* Journal of Clinical Pathology, 2019. **72**(10): p. 677-681.
4. Feng, M., et al., *Myosin light chain 9 promotes the proliferation, invasion, migration and angiogenesis of colorectal cancer cells by binding to Yes-associated protein 1 and regulating Hippo signaling.* Bioengineered, 2022. **13**(1): p. 96-106.
5. Deng, S., et al., *MYL9 expressed in cancer-associated fibroblasts regulate the immune microenvironment of colorectal cancer and promotes tumor progression in an autocrine manner.* Journal of Experimental & Clinical Cancer Research, 2023. **42**(1): p. 294.
6. Sellers, J.R., *Myosins: a diverse superfamily.* Biochimica et Biophysica Acta (BBA)-Molecular Cell Research, 2000. **1496**(1): p. 3-22.

7. Harrington, W.F. and M.E. Rodgers, *Myosin.* Annual review of biochemistry, 1984. **53**(1): p. 35-73.

8. Sun, J., et al., *Distinct roles of smooth muscle and non-muscle myosin light chain-mediated smooth muscle contraction.* Frontiers in Physiology, 2020. **11**: p. 593966.

9. Hirasawa, Y., et al., *Methylation status of genes upregulated by demethylating agent 5-aza-2'-deoxycytidine in hepatocellular carcinoma.* Oncology, 2007. **71**(1-2): p. 77-85.

10. Luo, X.-G., et al., *Histone methyltransferase SMYD3 promotes MRTF-A-mediated transactivation of MYL9 and migration of MCF-7 breast cancer cells.* Cancer letters, 2014. **344**(1): p. 129-137.

11. Zhao, S., W. Xiong, and K. Xu, *MiR-663a, regulated by lncRNA GAS5, contributes to osteosarcoma development through targeting MYL9.* Human & Experimental Toxicology, 2020. **39**(12): p. 1607-1618.

12. Walsh, M.P., et al., *Calcium-independent myosin light chain kinase of smooth muscle. Preparation by limited chymotryptic digestion of the calcium ion dependent enzyme, purification and characterization.* Biochemistry, 1982. **21**(8): p. 1919-1925.

13. Small, J.V. and A. SOBIESZEK, *Ca-Regulation of Mammalian Smooth Muscle Actomyosin via a Kinase-Phosphatase-Dependent Phosphorylation and Dephosphorylation of the 20000-Mr Light Chain of Myosin.* European Journal of Biochemistry, 1977. **76**(2): p. 521-530.

14. Badineni, M., et al., *Structure Elucidation and Identification of Novel Lead Molecules against Sulfur Import Protein cysA of Mycobacterium tuberculosis.* Current Protein and Peptide Science, 2023. **24**(7): p. 589-609.

15. Malkhed, V., et al., *Study of interactions between Mycobacterium tuberculosis proteins: SigK and anti-SigK.* Journal of molecular modeling, 2011. **17**(5): p. 1109-1119.

16. Malkhed, V., et al., *Identification of novel leads applying in silico studies for Mycobacterium multidrug resistant (MMR) protein.* Journal of Biomolecular Structure and Dynamics, 2014. **32**(12): p. 1889-1906.

17. Senior, A., et al., *AlphaFold: Using AI for scientific discovery.* DeepMind. Recuperado de: https://deepmind. com/blog/alphafold, 2018.

18. Varadi, M., et al., *AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models.* Nucleic acids research, 2022. **50**(D1): p. D439-D444.

19. Jumper, J., et al., *Applying and improving AlphaFold at CASP14.* Proteins: Structure, Function, and Bioinformatics, 2021. **89**(12): p. 1711-1721.

20. da Silva, G.M., et al., *Predicting Relative Populations of Protein Conformations without a Physics Engine Using AlphaFold 2.* bioRxiv, 2023.

21. Hooft, R.W., et al., *Errors in protein structures.* Nature, 1996. **381**(6580): p. 272-272.

22. Wiederstein, M. and M.J. Sippl, *ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins.* Nucleic acids research, 2007. **35**(suppl_2): p. W407-W410.

23. Ziemba, P., et al., *Using the PROSA method in offshore wind farm location problems.* Energies, 2017. **10**(11): p. 1755.

24. Prajapat, R., A. Marwal, and R. Gaur, *Recognition of Errors in the Refinement and Validation of Three-Dimensional Structures of AC1 Proteins of Begomovirus Strains by Using ProSA-Web.* Journal of Viruses, 2014. **2014**(1): p. 752656.

25. 攀苏, *Stereochemistry of polypeptide chain configurations.* J. mol. Biol, 1963. **7**: p. 95-99.

26. Pauling, L., R.B. Corey, and H.R. Branson, *The structure of proteins: two hydrogen-bonded helical configurations of the polypeptide chain.* Proceedings of the National Academy of Sciences, 1951. **37**(4): p. 205-211.

27. Elfmann, C. and J. Stülke, *PAE viewer: a webserver for the interactive visualization of the predicted aligned error for multimer structure predictions and crosslinks.* Nucleic acids research, 2023. **51**(W1): p. W404-W410.

28. Dundas, J., et al., *CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues.* Nucleic acids research, 2006. **34**(suppl_2): p. W116-W118.

29. Tian, W., et al., *CASTp 3.0: computed atlas of surface topography of proteins.* Nucleic acids research, 2018. **46**(W1): p. W363-W367.

30. Connolly, M.L., *Analytical molecular surface calculation.* Applied Crystallography, 1983. **16**(5): p. 548-558.

31. Connolly, M.L., *Solvent-accessible surfaces of proteins and nucleic acids.* Science, 1983. **221**(4612): p. 709-713.

32. Halgren, T., *New method for fast and accurate binding-site identification and analysis.* Chemical biology & drug design, 2007. **69**(2): p. 146-148.

33. Halgren, T.A., *Identifying and characterizing binding sites and assessing druggability.* Journal of chemical information and modeling, 2009. **49**(2): p. 377-389.

34. Lionta, E., et al., *Structure-based virtual screening for drug discovery: principles, applications and recent advances.* Current topics in medicinal chemistry, 2014. **14**(16): p. 1923-1938.

35. Sastry, M., et al., *Large-scale systematic analysis of 2D fingerprint methods and parameters to improve virtual screening enrichments.* Journal of chemical information and modeling, 2010. **50**(5): p. 771-784.

36. Nada, H., A. Elkamhawy, and K. Lee, *Identification of 1H-purine-2, 6-dione derivative as a potential SARS-CoV-2 main protease inhibitor: molecular docking, dynamic simulations, and energy calculations.* PeerJ, 2022. **10**: p. e14120.

37. Shelley, J.C., et al., *Epik: a software program for pK a prediction and protonation state generation for drug-like molecules.* Journal of computer-aided molecular design, 2007. **21**: p. 681-691.

38. Greenwood, J.R., et al., *Towards the comprehensive, rapid, and accurate prediction of the favorable tautomeric states of drug-like molecules in aqueous solution.* Journal of computer-aided molecular design, 2010. **24**(6): p. 591-604.

39. Jorgensen, W.L., D.S. Maxwell, and J. Tirado-Rives, *Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids.* Journal of the american chemical society, 1996. **118**(45): p. 11225-11236.

40. Madhavi Sastry, G., et al., *Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments.* Journal of computer-aided molecular design, 2013. **27**: p. 221-234.

41. Jorgensen, W.L. and J. Tirado-Rives, *The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin.* Journal of the American Chemical Society, 1988. **110**(6): p. 1657-1666.

42. Shivakumar, D., et al., *Prediction of absolute solvation free energies using molecular dynamics free energy perturbation and the OPLS force field.* Journal of chemical theory and computation, 2010. **6**(5): p. 1509-1519.

43. Friesner, R.A., et al., *Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy.* Journal of medicinal chemistry, 2004. **47**(7): p. 1739-1749.

44. Erickson, J.A., et al., *Lessons in molecular recognition: the effects of ligand and protein flexibility on molecular docking accuracy.* Journal of medicinal chemistry, 2004. **47**(1): p. 45-55.

45. Gilson, M.K. and H.-X. Zhou, *Calculation of protein-ligand binding affinities.* Annu. Rev. Biophys. Biomol. Struct., 2007. **36**(1): p. 21-42.

46. Morris, G.M. and M. Lim-Wilby, *Molecular docking.* Molecular modeling of proteins, 2008: p. 365-382.

47. Harrison, R.L. *Introduction to monte carlo simulation.* in *AIP conference proceedings.* 2010.

48. Friesner, R.A., et al., *Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein− ligand complexes.* Journal of medicinal chemistry, 2006. **49**(21): p. 6177-6196.

49. Halgren, T.A., et al., *Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening.* Journal of medicinal chemistry, 2004. **47**(7): p. 1750-1759.

50. Dixon, S.L., et al., *Medicinal Chemistry.*

51. Kollman, P.A., et al., *Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models.* Accounts of chemical research, 2000. **33**(12): p. 889-897.

52. Tsui, V. and D.A. Case, *Theory and applications of the generalized Born solvation model in macromolecular simulations.* Biopolymers: Original Research on Biomolecules, 2000. **56**(4): p. 275-291.

53. Srinivasan, J., et al., *Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate− DNA helices.* Journal of the American Chemical Society, 1998. **120**(37): p. 9401-9409.

54. Hou, T., et al., *Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations.* Journal of chemical information and modeling, 2011. **51**(1): p. 69-82.

55. Genheden, S. and U. Ryde, *The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities.* Expert opinion on drug discovery, 2015. **10**(5): p. 449-461.

56. Ausaf Ali, S., et al., *A review of methods available to estimate solvent-accessible surface areas of soluble proteins in the folded and unfolded states.* Current Protein and Peptide Science, 2014. **15**(5): p. 456-476.

57. Ashraf, N., et al., *Combined 3D-QSAR, molecular docking and dynamics simulations studies to model and design TTK inhibitors.* Frontiers in Chemistry, 2022. **10**: p. 1003816.

58. Durham, E., et al., *Solvent accessible surface area approximations for rapid and accurate protein structure prediction.* Journal of molecular modeling, 2009. **15**: p. 1093-1108.

59. Shrake, A. and J.A. Rupley, *Environment and exposure to solvent of protein atoms. Lysozyme and insulin.* Journal of molecular biology, 1973. **79**(2): p. 351-371.

60. Mostrag-Szlichtyng, A. and A. Worth, *Review of QSAR models and software tools for predicting biokinetic properties.* Luxembourg, 2010. **10**: p. 94537.

61. Ntie-Kang, F., *An in silico evaluation of the ADMET profile of the StreptomeDB database.* Springerplus, 2013. **2**: p. 1-11.

62. Zhu, Y., et al., *ADME/toxicity prediction and antitumor activity of novel nitrogenous heterocyclic compounds designed by computer targeting of alkylglycerone phosphate synthase.* Oncology letters, 2018. **16**(2): p. 1431-1438.

63. Madhavaram, M., et al., *High-throughput virtual screening, ADME analysis, and estimation of MM/GBSA binding-free energies of azoles as potential inhibitors of Mycobacterium tuberculosis H37Rv.* Journal of Receptors and Signal Transduction, 2019. **39**(4): p. 312-320.

64. Guengerich, F.P., *Cytochrome p450 and chemical toxicology.* Chemical research in toxicology, 2008. **21**(1): p. 70-83.

65. Pelkonen, O., et al., *Inhibition and induction of human cytochrome P450 (CYP) enzymes.* Xenobiotica, 1998. **28**(12): p. 1203-1253.