Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Integrated Machine Learning, Molecular Docking, and Molecular Dynamics Simulations Design Novel Peptide Inhibitors of Rabies Virus Glycoprotein

Harshada Sawant

sawant.harshada10@gmail.com

Haffkine Institute for Training, Research and Testing

Ashish Asrondkar

Haffkine Institute for Training, Research and Testing

Vrushali Patil

Haffkine Institute for Training, Research and Testing

Research Article

Keywords: Rabies, Glycoprotein, α-bungarotoxin, neurotoxic, peptides, ML-DL hybrid framework, molecular dynamic simulations

Posted Date: August 3rd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10274307/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

TITLE: Integrated Machine Learning, Molecular Docking, and Molecular Dynamics Simulations Design Novel Peptide Inhibitors of Rabies Virus Glycoprotein

Harshada Sawant,* Ashish Asrondkar, Vrushali Patil

Department of Chemotherapy, Haffkine Institute for Training, Research and Testing, Parel Mumbai Maharashtra, India 400012

Corresponding Author: Harshada Sawant

Email id: sawant.harshada10@gmail.com

Department of Chemotherapy, Haffkine Institute for Training, Research and Testing, Parel Mumbai Maharashtra, India 400012

ABSTRACT: Rabies is the most lethal viral zoonotic diseases worldwide. The rabies virus glycoprotein (RABV-G) which mediates host cell recognition and viral entry; it represents a therapeutic target for the development of novel anti-viral agents. In the present study, a computational peptide design strategy was employed to design peptide inhibitors against RABV-G. Peptide sequences were designed from the receptor binding region of α-bungarotoxin through mutation and optimisation while eliminating neurotoxic properties. A hybrid machine learning-deep framework was used for toxicity screening. The selected non-toxic peptides were analysed through framework of different tools for Physiochemical characterization, antiviral activity prediction, molecular docking, binding affinity and molecular dynamic simulation were performed to analyse peptide RABV-G interactions. A three-dimensional model of RABV-G was generated using SWISS MODEL and validated through Ramchandran plot analysis with 92.5% residues within favoured regions. Binding pocket analysis identified key residues involved in peptide binding including Glu272, Glu279., Asp283, Lys276 and Arg297. Among the evaluated candidates, MKMKCHTTATMKIV (P40) shown the most favourable docking with a HADDOCK scores -90 ± 3.8, and binding free energy of -10.3kcal/mol, followed by CLDLGYTIVCHTTATSPI (P3), MKIVCHTTATSPMKV (P21), and MKIVCHTTMKSPMK (P25). Molecular dynamic simulations (100ns) analysed that P40 and P25 formed highly stable complexes with RABV-G, exhibiting low RMSD and RMSF fluctuation with intermolecular interaction, and favourable free energy landscapes and stable conformational through simulation period. This study design shows the α-bungarotoxin derived peptides can target RABV-G and P440 can be promising lead candidate for development of peptide-based drug delivery.

KEYWORDS: Rabies, Glycoprotein, α-bungarotoxin, neurotoxic, peptides, ML-DL hybrid framework, molecular dynamic simulations

INTRODUCTION: Rabies is an acute viral disease which causes encephalitis in almost all warm-blooded animals. It is one of the most dreaded zoonotic disease due to the traumatic and painful experience during the course of the infection.(1) Human rabies is mainly transmitted through a rabid dog bite in the developing world, of which 94-98% of deaths are due to canine rabies.(2) World-wide human mortality from rabies is estimated to be 55000 deaths per year with 56% of these deaths estimated to occur in Asia and 44% in Africa.(3) In India alone, 18500 people die and some 70,0000 people take rabies prophylaxis each year following exposure to rabid animals. (4) The Rabies is caused due to zoonotic, neurotropic non-segmented RNA viruses which also known as RABV virus.(5) RABV belongs to the family of Rhabdoviridiae, genus of Lyssavirus, order Mononegavirales. It bears stark resemblance to the Vesicular Somatitis virus (VSV). (6) It is bullet shaped viron with a plasma membrane covered with homotrimers of type 1 of glycoproteins (G).(7)

RABV of *Lyssaviruses* consist of five viral proteins encoding Nucleoprotein(N), Glycoprotein (G) Phosphoprotein (P) Matrix Protein (M) and large RNA polymerase protein (L).(8) Owing to the location of G protein on the viral surface, it is considered the major determinant of tissue tropism.(9) The RV exhibits a broad host spectrum, highlighting the importance of G protein in interacting with multiple host receptors. The RABV-G protein is a type I membrane glycoprotein which is translated on membrane-bound ribosomes and inserted co-translational into the endoplasmic reticulum (ER) in an unfolded form.(10,11) The folding of transmembrane proteins occurs in three topologically and biochemically distinct environments: the ER lumen, the ER membrane, and the cytosol. The structural organization of RABV-G protein indicated the constitution of three domains; ectodomain, transmembrane domain (TMD), and the cytoplasmic domain which can fold independently of each other.(12,13) Cellular receptors are regarded as the primary pathway through which viruses can gain access to the host. The RABV interacts with a wide range of receptors by which it enters cells of different host species. Successful infection occurs only when a receptor can initiate the full viral life cycle from cell recognition to release the genome into the cell for replication and protein synthesis.(14,15) Elucidating the preferences of RABV in binding and attaching to certain types of host receptors is a point of interest which may provide insights into how entry mechanisms may be targeted therapeutically.(16) RABV-G protein is also primary antigen responsible for inducing virus neutralizing antibodies, due to which it make key target receptor for vaccine development and designing antiviral strategies. (17)

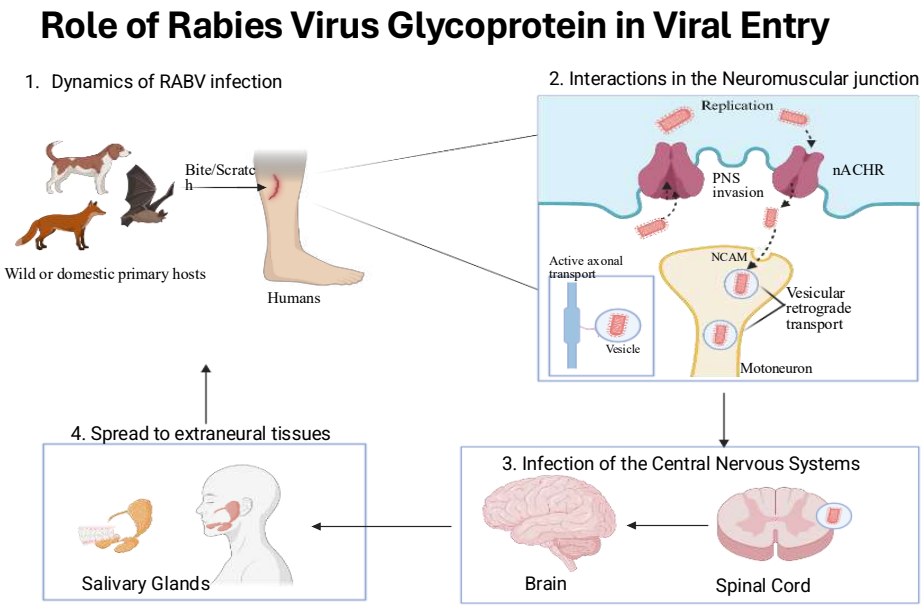


Fig 1: RABV-G entry in Host cell.

Peptides are ideally suited to mimic natural ligands and thereby function in an antagonistic or agonistic way. Furthermore, they are able to physiologically disrupt functional complexes due to their small size and specific binding properties. (18–20)

Peptides have moved into the focus of drug discovery, especially because of their small size, their structural diversity, their high degree of specificity towards a given target, their easy selection from large natural or artificial peptide libraries, and their easy chemical synthesis. The in vivo pharmacokinetics of peptides, like that of other drugs, depends on a number of factors such as absorption, distribution, metabolism and excretion collectively referred to as "ADME".(21–23)

Alpha-bungarotoxin (α- bungarotoxin) was one of the major proteins isolated from the crude venom of the elapid snake *Bungarus multicinctus*. It belonged to the class of neurotoxins and was well known for causing respiratory failure and flaccid paralysis through its pharmacological action at the neuromuscular junction. α- bungarotoxin was, in fact, a postsynaptic neurotoxin that prevented neurotransmission by selectively binding to nicotinic acetylcholine receptors (AChR) on the postsynaptic membrane. (24,25)

The rabies virus and alpha-bungarotoxin (α- bungarotoxin) bind competitively to the same site on the nicotinic acetylcholine receptor (AChR).(26) The rabies virus binds to AChR through a

fragment of the virus glycoprotein, which contains structural features shared with neurotoxins such as α- bungarotoxin.(27) In fact, at picomolar concentrations, the glycoprotein competes with α- bungarotoxin for binding to AChR.

The structures of residues 175-203 of the rabies virus glycoprotein involved in this competition are like neurotoxin loop 2.(28,29) Synthetic peptides mimicking this neurotoxic loop 2 as well as the glycoprotein of the rabies virus can inhibit the receptor binding of α- bungarotoxin with inhibition values not too different from those of nicotine and D-tubocurarine. (30,31)

In addition, it has been demonstrated that monoclonal antibodies raised against residues 190–203 of the rabies virus glycoprotein inhibit both the rabies virus glycoprotein and α-bungarotoxin from binding to the AChR.(32) This suggests that the rabies virus binds to the nicotinic AChR and that at least in part a portion of the glycoprotein, which resembles neurotoxic loop 2, mediates this binding. (33)

According to the research, rabies virus is binding to synthetic peptide (Tal 173-204), which is produced from AChR α-subunit in some specific manner.(34) Synthetic peptides of neurotoxin and glycoprotein both can inhibit this interaction. However, it was observed that the affinities for α- bungarotoxin binding were higher than the affinities of rabies virus binding to the receptor.(26,35)

Based on this study, the short peptides were designed by mimicking the bungarotoxin amino acids which interact with the receptor, RAV-G, by eliminating the toxin and its effect of neurotoxicity. These engineered peptides might bind to the RAV-G receptor thereby disrupting virus host interactions. Investigating these peptide motifs through computational modelling, molecular docking, and experimental binding studies can provide insights into their ability to block glycoprotein–receptor interactions. Ultimately, this strategy may contribute to the development of novel peptide-based antiviral agents that target early stages of rabies virus infections

MATERIALS & METHODOLOGY.

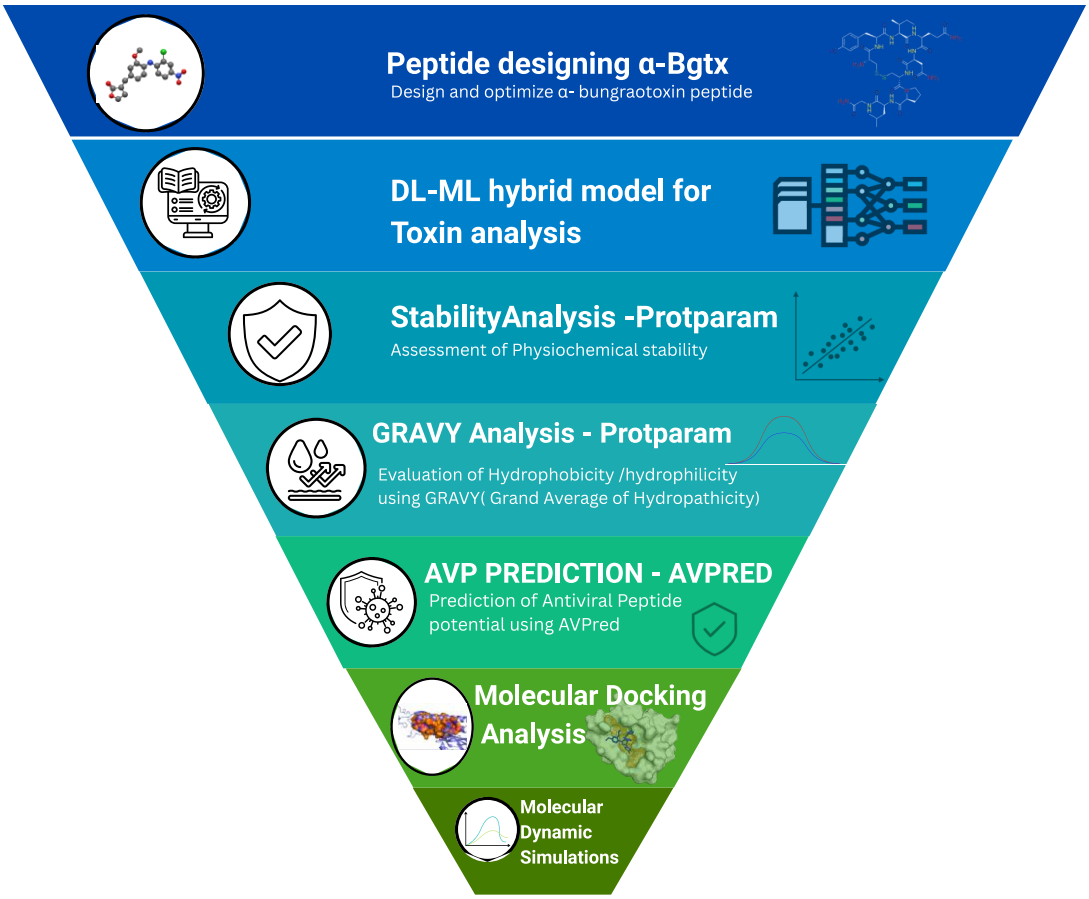


Fig 2 : Flowchart for the peptide designing

1. Target Selection and Homology Modelling

The *Rabies lyssavirus* glycoprotein (RABV-G), the major viral envelope protein responsible for receptor recognition and host-cell entry, was selected as the primary molecular target. The full-length amino acid sequence of RABV-G was retrieved from UniProt and used to generate a 3D structural model using SWISS-MODEL (https://swissmodel.expasy.org/). Template search was performed within the SWISS-MODEL Template Library, and models were ranked based on GMQE, QMEAN, and global stereochemical quality. The top-scoring model was selected for downstream analyses. (36,37)

2. Model Validation and Pre-processing

Structural validation of the modelled glycoprotein was carried out using SAVES v6.1 (https://saves.mbi.ucla.edu/), employing PROCHECK, VERIFY3D, and ERRAT. Ramachandran plots ensured >90% residues were within favoured/allowed regions.(38–42) The structure was cleaned in PyMOL (https://pymol.org/) by removing water molecules and correcting termini, followed by energy minimisation to prepare the receptor for docking.(43)

3. Design of Peptide Analogues

1. Peptide Design and Sequence Identification:

The amino acid sequence of α-bungarotoxin; GenBank: CAA63045.1 (IVCHTTATSPISAVTCPPGENLCYRKMWCDAFCSSRGKVVELGCAATCPSKKPY EEVTCCSTDKCNPHPKQRPG) was used as the primary template for peptide design. To identify potential functional peptide sequences, the sequence was analyzed to locate regions responsible for receptor binding. Particular importance was placed on sequences enriched with aromatic (tyr, Trp), charged (Arg, Lys), and Polar residues, which contribute to protein-receptor interactions. The sequence was fragmented into smaller overlapping peptides using a sliding window approach.(44) Peptides of 15-20 amino acids were generated to retain functional binding motifs. The further optimization was done through systematic point mutategenesis to form a comprehensive library of variant peptides. Each amino acid position within the parent sequences was sequentially substituted with the remaining 19 standard amino acids, excluding the 19 residues. And these variants were subsequently screened for physiochemical, toxicity and AVP properties while maintaining binding affinity and structural stability and specificity.(45)

2. Toxicity Prediction Using Machine Learning and Deep Learning

Peptide toxicity was predicted using an integrated computational framework combining machine learning(ML) and deep learning (DL) approaches. Each peptide sequence was transformed into a numerical feature using physiochemical descriptors, including hydrophobicity, hydropathicity, hydrophilicity, net charge, molecular weight and amino acid composition. A curated dataset was pseudo-labeled using domain-informed thresholds, where peptides with net charge>6.5 and Boman index>2.5 were classified as toxic. The dataset was split into training (80%) and testing (20%) sets. Baseline toxicity predicition using machine learning models such as Random forest, support vector machine and K-Nearest Neighbours. To captur complex and non-linear relationships, a deep neural network (DNN) with fully connected layers, ReLU activation, batch normalization and dropout regularization wqas implemented and optimised using the Adam optimizer with binary cross-entropy loss. Additionally, an unsupervised autoencoder model was employed for anomaly detection, where peptides were high reconstruction error were identified as potentially toxic. (46–48)

The 201 shortlisted peptides were further subjected to toxicity prediction to refine the candidate pool using a hybrid deep learning framework integrating both supervised and unsupervised learning approaches. A deep feedforward neural network (DNN) was employed for supervised classification of peptide toxicity based on physicochemical and molecular descriptors. In parallel, an unsupervised autoencoder model was utilized to identify potential toxic outliers based on reconstruction error, thereby improving detection sensitivity in datasets with limited or incomplete toxicity annotations (49)

This dual-model strategy enhanced the robustness and generalizability of the screening process by capturing complex and non-linear toxicological patterns beyond conventional descriptor-based thresholds. As a result, 163 peptides were predicted to be potentially toxic and excluded from further analysis, yielding a final set of 40 non-toxic peptides (Supplementary Material 3).

FASTA sequences were generated using the ReadSeq Sequence Format Converter (http://avermitilis.ls.kitasato-u.ac.jp/readseq.cgi), and PDB structures were exported for docking.

4. Physicochemical Property Prediction (ProtParam)

The physicochemical characteristics of each peptide sequence were computed using the ProtParam tool available on the ExPASy server (ProtParam; https://web.expasy.org/protparam/). The analysis generated the following descriptors: amino acid composition, molecular weight (Da), theoretical isoelectric point (pI), total numbers of positively and negatively charged residues, instability index, aliphatic index, grand average of hydropathicity (GRAVY), extinction coefficient, and estimated half-life in *Homo sapiens* (mammalian reticulocytes, in vitro), *Saccharomyces cerevisiae* (yeast, in vivo), and *Escherichia coli* (bacteria, in vivo) systems. Additionally, the chemical formula and total atomic composition for each peptide were calculated. These physicochemical parameters were subsequently used to evaluate peptide solubility, thermostability, and overall biochemical behaviour relevant to peptide functionality and biological activity.(50,51)

6. Antiviral Activity Prediction (AVP-IC_{50}Pred)

The antiviral activity of each peptide was predicted using AVP-IC50Pred (http://crdd.osdd.net/servers/ic50avp/). Peptide sequences were analysed using the General Predictor regression model to estimate their inhibitory potential. The tool generated the following parameters: predicted IC_{50} values (μM), regression-based antiviral activity scores, and rank-ordering of peptides based on predicted potency.(52)

Lower IC_{50} values indicated stronger predicted antiviral efficacy, thereby enabling prioritization of peptide candidates with high therapeutic potential.

7. Binding Site Identification on Glycoprotein

The major binding pockets of the Glycoprotein were identified to guide subsequent molecular docking studies. Binding site prediction was performed using the DoGSiteScorer algorithm integrated into the ProteinsPlus web server (https://proteins.plus). This automated tool analyzes the geometric and physicochemical properties of the protein surface to detect potential druggable pockets. The top-ranked pockets, based on volume, depth, and druggability score, were selected for further molecular docking and dynamics studies. (53)

Molecular Docking:

Protein Peptide molecular docking was performed using the HADDOCK (High Ambiguity Driven Biomolecular Docking) server to identify the intermolecular interactions and binding mode between the receptor and ligand protein/peptides.(54) Prior to docking the receptor was pre-processed by removing water molecules, heteroatoms and co-crystallized ligand to obtain a receptor structure and Peptide structure was modelled using PEPfold3 suitable for interacting studies. Missing atoms or residues were corrected and energy minimized before docking.

The prepared receptor and peptide structure were submitted to the HADDOCK server for protein-peptide docking. Active and passive interaction residues were defined basis of DoGsitescore for pocket residues. Docking simulation was runed using standard HADDOCK protocol consisting of rigid-body energy minimization , annealing refinement, and explicit solvent refinement in water. Multiple conformations of docked complex were generated and clustered on the basis of structural similarity using Fraction of Common contacts (FCC) clustering algorithms. The resulting complexes were ranked according to the HADDOCK score which integrates Van Der Waals Energy, Electrostatic interactions, Desolvation energy,

Restraint violation energy and buried surface area to estimate docking quality and Structural compatibility.

The top-ranked docked complexes showing the most favourable HADDOCK score and interaction profile was analysed using PRODIGY to predict binding affinity and thermodynamic stability.(55–57)

Molecular Dynamic Simulations

Molecular Dynamic Simulations (MDS) was performed using the Desmond module(Academic Version 2024) to evaluated the structural stability, conformational behaviour and dynamic interactions of the protein-peptide complex under physiological conditions. The highest scoring complex was used as the initial structure for the simulation analysis. Initially, the system was prepared using Protein Preparation Wizard, wherein missing hydrogen atoms were added, bond orders were assigned and missing side chains or loops were added when necessary. Water molecules were removed beyond 5.0Å and the amino acid residues were optimised at pH 7.5. The prepared complex was then embedded in cubic simulation box and solvated using TIP3P(Transferrable intermolecular Potential with 3 Points) , 3-site water model with an appropriate buffer distance around the complex. Counter ions were added to neutralize the system charge and salt concentration (0.15M NaCl) was maintained to mimic biological conditions. Energy minimization was of solvated system was carried out followed by system equilibration. The molecular docking simulation was carried under an NPT ensemble at 300K temperature and 1atm pressure using OPLS force field for 100ns time. Temperature was maintained by Nose-Hoover thermostat and Pressure was maintained by Martyna-Tobias-Klein barostat respectively. (53)

Post Molecular Dynamic Simulations: Trajectory analysis was evaluated for dynamic stability and structural fluctuations of the complex through the simulation period. The parameters root mean square deviation (RMSD), root mean square fluctuation (RMSF), Radius of gyration (Rg), solvent accessible surface area (SASA), hydrogen bonding interactions were done using Simulations event analysis and the graphs were ploltted. Prinicpal component analysis (PCA), Free energy landscape analysis (FEL) and Dynamic Cross-Correaltion Matrix (DCCM) was plotted using R-studio and its libraries.

RESULT AND DISCUSSION:

Protein molecular modelling: Gene Bank accession number for Rabies *Lyssavirus* Glycoprotein is LOCUS: AAC34683 ACCESSION NO:AAC34683 VERSION AAC34683.1 is used for protein molecular modelling.

Primary amino acid sequence of *Lyssavirus* Glycoprotein is as follows.

MVPQVLLFVP LLGFSLCFGK FPIYTIPDKL GPWSPIDIHH LSCPNNLVVE DEGCTNLSEFSYMELKVGYI SAIKVNGFTC TGVVTEAETY TNFVGYVTTT FKRKHFRPTP DACRAAYNWKMAGDPRYEES LHNPYPDYHW LRTVRTTKES LIIISPSVTD LDPYDKSLHS RVFPGGKCSG ITVSSTYCST NHDYTIWMPE NPRPRTPCDI FTNSRGKRAS KGNKTCGFVD ERGLYKSLKG ACRLKLCGVL GLRLMDGTWV AMQTSDETKW CPPDQLVNLH DFRSDEIEHL VVEELVKKREECLDALESIM TTKSVSFRRL SHLRKLVPGF GKAYTIFNKT LMEADAHYKS VRTWNEIIPSKGCLKVGGRC HPHVNGVFFN GIILGPDGHV LIPEMQSSLL QQHMELLKSS VIPLMHPLAD PSTVFKEGDE AEDFVEVHLP DVYKQISGVD LGLPNWGKYV LMTAGAMIGL VLIFSLMTWCRRANRPESKQ RSFGGTGRNV SVTSQSGKVI PSWESYKSGG EIRL [21]

The 3D model of Protein was done Using Swiss Model Server. The Template 6lgx.1. A is selected for modeling protein with the sequence identity 100% and coverage 100% for modeling.(13) The SWISS-MODEL template library (SMTL version 2020-04-08, PDB release 2020-04-03) is searched with BLAST for phylogenetically matched structures with the target sequence.

Modelled structure obtained from SWISSMODEL server has -0.88 QMEAN score. QMEAN value is a linear combination of four statistical relevant terms. It provides a Z score to relate it to high resolution X-ray structures with similar size. Higher Z score is related to more favourable model. Ramachandran Plots (fig 2) are drawn for this model by using web servers of ProCheck by SAVES 6.O. It is calculated that, modelled structure has 92.5% residues in most favoured regions, 6.6% in additional allowed regions, 0.3% in generously allowed regions and 0.6% in disallowed regions. (58)

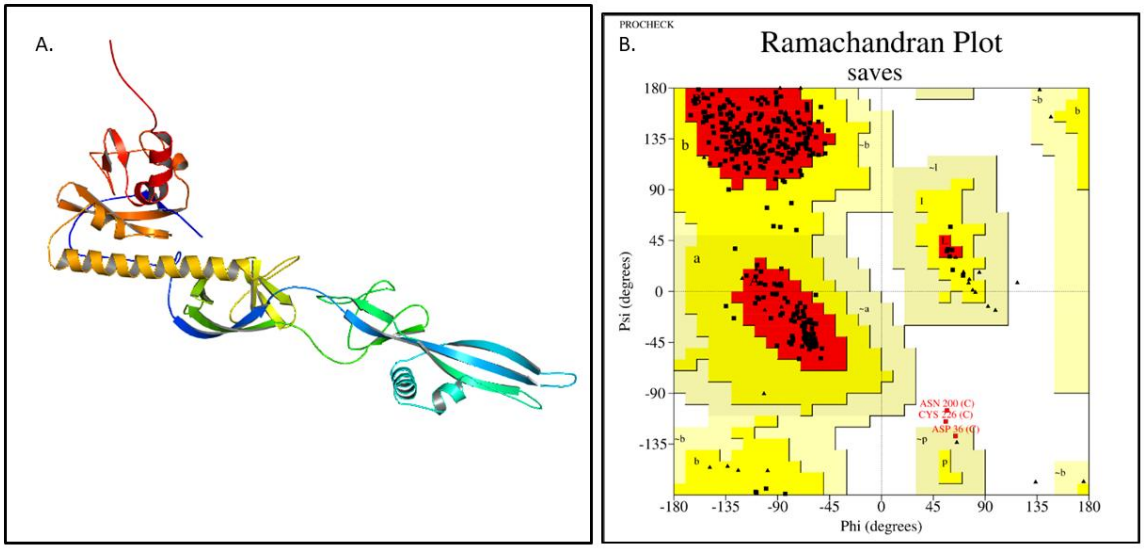


Fig 3: Structural Validation of Rabies Glycoprotein model. A. Predicted three-dimensional structure of RABV-G represented in Ribbion diagram. B. Ramchandran plot of the modelled structure.

Design Strategy and Mutation Workflow for the Peptide Inhibitor Library

The peptide inhibitor library was developed through a structured mutation–selection approach using the parent inhibitor sequence (inh) as the foundational template. Amino acid substitutions were guided by their biochemical classification to ensure systematic and chemically coherent modifications. For this purpose, residues were grouped into six major classes based on shared physicochemical properties: acidic (Asp, Glu), basic (Arg, His, Lys), hydrophobic/aliphatic (Ala, Ile, Leu, Met, Val), aromatic (Phe, Trp, Tyr), polar uncharged (Asn, Cys, Gln, Ser, Thr), and unique (Gly, Pro).

Single-point mutations were introduced sequentially across the peptide backbone, with each residue replaced only by alternatives from its respective chemical group. This ensured that hydrophobic residues such as leucine, valine, and isoleucine were substituted exclusively within the aliphatic class, whereas polar or charged residues—including lysine, cysteine, histidine, and threonine—were interchanged among chemically compatible side-chain types. This group-restricted substitution strategy minimized structural disruption, preserved electrostatic balance, and prevented the combinatorial explosion associated with unrestricted mutational scanning.

All single-residue variants were evaluated through molecular docking to assess their binding affinity toward the selected molecular target. Mutants exhibiting improved binding compared to the parent sequence were retained and iteratively combined to produce higher-order derivatives containing two to four coordinated substitutions. Each generation of multi-point mutants underwent additional docking-based screening, and only those demonstrating progressively enhanced interaction energies were advanced.

Through this iterative cycle of mutation, computational evaluation, and rational recombination, a final set of 130 optimized peptide variants was generated. These sequences represent the refined inhibitor library enriched for improved predicted affinity and structural compatibility with the target protein.

ML of Toxicity Prediction:

Machine Learning (ML) and Deep Learning (DL) methods were integrated to predict peptide toxicity analysis using physiochemical descriptors and sequence based information. The conventional ML models including Random Forest (RF), Support Vector Machine (SVM), and K-Nearest Neighbors (KNN) were trained using peptide physiochemical properties such as hydrophobicity, hydropathicity, hydrophilicity, charge, and molecular weight. Among these, RF and SVM demonstrated superior predictive performance with an approximately 80% of accuracy whereas KNN showed comparatively lower accuracy of 73%. Receiver operating characterstics (ROC) analysis yielded an accuracy of Area under the Curve (AUC)of 0.76.

Prinicpal component analysis (PCA) demonstrated partial clustering of toxic and non-toxic peptides with considerable overlapping, indicating that physiochemical descriptors alone are insufficient and that peptide toxicity is influenced by complex and non-linear molecular interactions. To improve predictive performance, a hybrid Deep learning-machine learning (DL-ML) framework integrating Convolution Neural Networks (CNNs) with physiochemical descriptors was integrated. The model demonstrated enhanced capacity compared to conventional ML approaches by effective study of local sequence motifs and positional amino acid ranking with peptide toxicity. Training and Validation curves indicated successful feature learning despite fluctuations in accuracy and loss. Minimal divergence between training and validation performance suggested acceptable generalization with limited overfitting.

Based on the DL-ML toxicity prediction results, peptides were ranked according to their predicted toxicity profiles, and the top 40 peptides exhibiting the non-toxic nature and then were subsequent downstream analysis.(59–61)

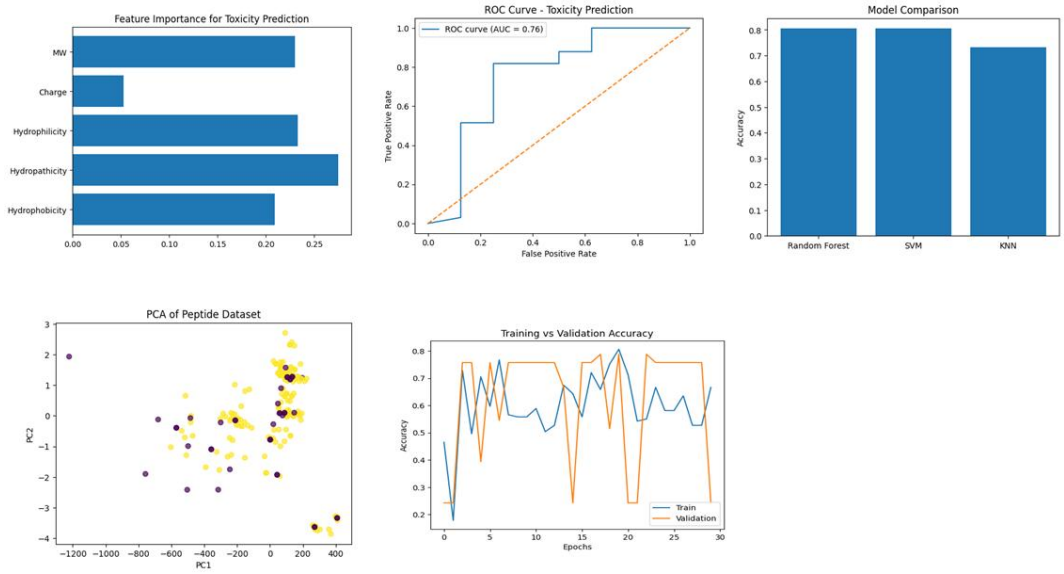


Fig 4 : Machine Learning-Deep Learning Hybrid based toxicity prediction analysis of designed peptides: Feature importance, ROC analysis, model comparison, PCA clustering, Training-validation accuracy.

Table 1: Toxin Prediction properties

Code	Sequence	Prediction	Hydrophobicity	Steric hindrance	Side bulk	Hydropathicity	Amphipathicity	Hydrophilicity
P3	CLDLGYTIVCHTTATSPI	Non-Toxin	0.08	0.56	0.56	0.78	0.08	-0.68
P21	MKIVCHTTATSPMKV	Non-Toxin	-0.06	0.58	0.58	0.37	0.59	-0.29
P25	MKIVCHTTMKSPMK	Non-Toxin	-0.16	0.6	0.6	-0.13	0.89	-0.01
P40	MKMKCHTTATMKIV	Non-Toxin	-0.14	0.61	0.61	0.12	0.89	-0.1

RVG-29	YTIWMPENPRPGTPCDIFTNSRGKRASNG	Non-Toxin	-0.26	0.61	0.61	-1.05	0.42	0.12

Physiochemical Properties

A total of forty peptide sequences were compiled and examined through a series of *in silico* evaluations that included physicochemical profiling, toxicity prediction, and assessments of intrinsic stability. These peptides differed considerably in their lengths, molecular masses, and residue compositions, which in turn produced a broad distribution of physicochemical behaviours. Their theoretical pI values spanned from 3.8 to 10.3, covering sequences expected to behave as either acidic or basic under physiological conditions and thus showing distinct solubility and interaction tendencies. (Table 2)

When instability indices were analysed, most peptides fell within the stable category (index < 40). This group comprised P1, P4, P5, P7–P9, P11, P14, P17, P20, P21, P26–P30, and P31–P40. A smaller subset—P2, P3, P6, P12, P13, P15, P16, P18, P19, P22, P24, and P25—displayed instability values above the accepted stability threshold, indicating a greater likelihood of rapid degradation. Such sequences can still be rendered suitable for biological applications through well-established stabilisation approaches, including N-terminal acylation or formylation, C-terminal amidation, incorporation of D-amino acids or non-canonical residues, cyclisation or other backbone modifications, covalent crosslinking, nanoparticle-based delivery formats, or strategic increases in molecular mass.

The aliphatic index, often used as an indicator of thermostability, varied widely. Notably, P2, P3, P4, P8, P9, P28, P29, and P30 exhibited values exceeding 100, suggesting enhanced thermal resilience. GRAVY values showed an equally broad distribution, from strongly negative to moderately positive. Peptides such as P14, P22, P25, and P38 were more hydrophilic, whereas P4, P8, and P9 were distinctly hydrophobic, implying a greater tendency for membrane association.

Predicted half-life values differed across mammalian, yeast, and *E. coli* systems. Several peptides—particularly S14, S20, S21, S31, S35, and S36—were projected to persist longer in

multiple biological contexts, indicating a higher probability of retaining functional activity during downstream applications.

Across the dataset, the extinction coefficient was calculated to be 4720 M^{-1} cm^{-1}, providing a useful parameter for quantitative studies involving peptide–protein or peptide–ligand interactions under solution conditions.

Antiviral potential assessed using AVP-IC_{50}Pred revealed IC_{50} values ranging from 37.8 to 78.1 μM, suggesting moderate inhibitory activity across the dataset. Notably, P40 (MKMKCHTTATMKIV) and P20 (MKIVCHTTATMKIV) demonstrated the lowest predicted IC_{50} values, indicating comparatively stronger antiviral potential than the remaining sequences. (Table 3)

Table 2: Physiochemical Properties of the designed peptides

Code	Sequence	No. of Amino Acids	Theoretical pI	Charged Residues (−)	Charged Residues (+)	Molecular Formula	Total Atoms	Estimated Half-life*	Instability Index	Stability	Aliphatic Index	GRAVY
P3	CLDLGYTIVCHTTATSPI	18	5.08	1	0	$C_{83}H_{134}N_{20}O_{27}S_2$	266	Yeast, *E. coli*	44.33	Unstable	108.33	0.778
P21	MKIVCHTTATSPMKV	15	9.31	0	2	$C_{70}H_{123}N_{19}O_{20}S_3$	235	Mammalian reticulocytes, Yeast, *E. coli*	33.33	Stable	71.33	0.367
P25	MKIVCHTTMKSPMK	14	9.79	0	3	$C_{69}H_{123}N_{19}O_{18}S_4$	233	Mammalian reticulocytes, Yeast, *E. coli*	41.06	Unstable	48.57	−0.129

P40	MKMKCHTTATMKIV	14	9.79	0	3	$C_{68}H_{123}N_{19}O_{18}S_4$	232	Mammalian reticulocytes, Yeast, *E. coli*	38.34	Stable	55.71	0.121
RVG_29	YTIWMPENPRPGTPCDIFTNSRGKRASNG	29	9.31	2	4	$C_{141}H_{217}N_{43}O_{43}S_2$	446	Mammalian reticulocytes, Yeast, *E. coli*	41.86	Unstable	30.34	-1.048

Table 3: Predicted IC_{50}

Code	Sequence	AVP-IC_{50}	
		Length	IC_50 (μM)
P3	CLDLGYTIVCHTTATSPI	20	47.71
P21	MKIVCHTTATSPMKV	17	45.7
P25	MKIVCHTTMKSPMK	16	46.41
P40	MKMKCHTTATMKIV	16	37.85
RVG_29	YTIWMPENPRPGTPCDIFTNSRGKRASNG	29	38.23

Structure–Activity Relationship (SAR) of Peptides in Relation to Predicted IC_{50} Values

Analysis of sequence composition, length variation, and physicochemical profiles revealed distinct structural determinants associated with the antiviral activity of the peptides, as reflected by their predicted IC_{50} values. Peptides exhibiting lower IC_{50} values (e.g., P20: MKIVCHTTATMKIV, P40: MKMKCHTTATMKIV) shared several conserved characteristics. These included (i) the presence of multiple basic residues (M, K), which likely enhance electrostatic interactions with viral surface components; (ii) the occurrence of Cys–His–Thr (CHT/CHTT)-containing motifs, which may contribute to structural rigidity and facilitate target-specific binding; and (iii) overall intermediate to high isoelectric points, supporting their ability to interact with negatively charged viral or host membranes.

Peptides with moderately higher IC_{50} values generally lacked one or more of these structural features. For example, sequences with reduced lysine content or shorter chain lengths (e.g., P7, P8) displayed weaker predicted antiviral potency, suggesting that reduced molecular surface area and fewer potential interaction sites diminish binding efficiency. Similarly, peptides with more acidic residues or lower isoelectric points showed relatively higher IC_{50} values, indicating that charge distribution plays a key role in antiviral activity.

Overall, the SAR analysis indicates that basic, moderately hydrophobic peptides containing extended CHT/CHTT-based motifs and 14–16 amino acid lengths exhibit the most favorable predicted antiviral activities. These trends highlight key structural elements governing peptide–target interactions and provide a rational basis for designing optimized antiviral peptide candidates.

Binding Pocket Analysis:

The binding pocket analysis was performed to identify the key amino acid residues involved in the active sites of modelled glycoprotein using computational prediction methods. For the glycoprotein molecule, the predicted binding pocket volume and surface area were 1777.45 Å and 2107.68 Å. The pocket contained 356 amino acids, 50 hydrogen bond donors and 113 acceptors, with a hydrophobicity ratio of 0.39. (Table 4, Fig5)

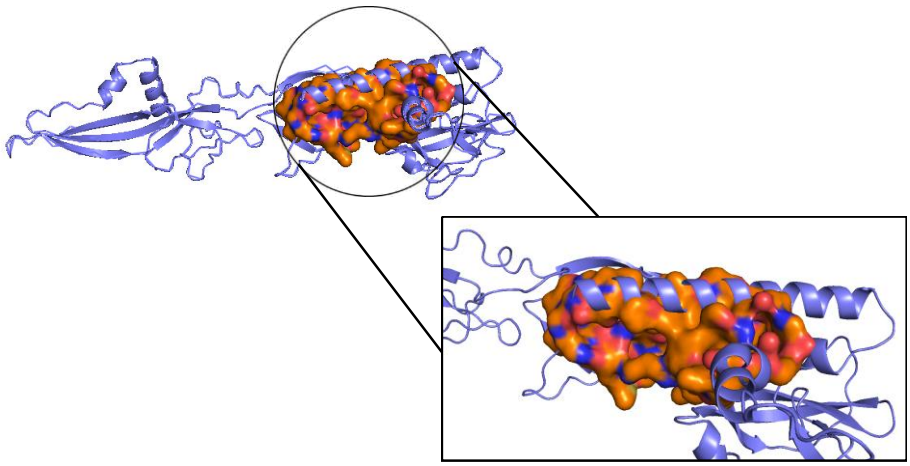


Fig 5: Predicted binding pocket of the modelled Rabies virus glycoprotein (RABV-G) shown as an orange surface (red indicating polar regions) mapped onto the protein cartoon (blue). Inset shows a zoomed view of the pocket.

Table 4: The amino acid residue present in the binding site of RABV-G

Sr. No.	Protein Name	Binding Pockets
1.	Glycoprotein	Lys 5, Phe6, Pro7, Ile8, Tyr9, Thr10, Cys28, Pro29, Asn30, Asn31, Leu32, Val33, Glu35, Tyr47, Glu49,Leu50, Phe207, Val208, Asp209, Arg211, Gly212, Leu213, Tyr214, Lys215, Ser216, Leu217, Lys218, Ala220, Leu233, Met234, Asp235, Gly236, Thr237, Trp238, Val239, Asn257, Leu258, His259, Asp260, Ser263, Asp264, Ile266, Glu267, Val270, Val271, Glu273, Leu274, Val275, Lys277, Arg278, Glu280, Cys282, Ser300, His301, Leu302, Arg303, Leu305, Val306, Pro307, Trp333, Pro372, Glu373, Ser376, Ser377

Molecular docking

A comprehensive in silico screening of 40 peptide sequences were performed using HADDOCK 2.4 and HDOCK, enabling a cross-validated assessment of binding affinity, interface complementarity, and energetic stability. The evaluation primarily focused on HADDOCK scores, electrostatic and van der Waals energy components, buried surface area (BSA), Z-scores, and HDOCK global docking scores.(57,62,63)

HADDOCK analysis revealed significant variations in binding energies among the peptides, with MKMKCHTTATMKIV (P40) emerging as the top-performing ligand, evidenced by its lowest HADDOCK score of −90.0 ± 3.8. The strong negative electrostatic energy (−411.1 kcal/mol) and large BSA of 1340.1 Å² indicate robust intermolecular attraction and an extensive interaction interface with the viral receptor. These properties suggest a highly stable and tightly bound peptide–protein complex. Similarly, CLDLGYTIVCHTTATSPI (P3) demonstrated a highly competitive HADDOCK score of −86.3 ± 4.5, coupled with the highest BSA (1645.9 ± 44.1 Å²) among all peptides, highlighting superior shape complementarity and surface engagement. The corresponding Z-score of −2.0, indicative of well-defined docking clusters, strengthens the confidence in its predicted binding pose.

Secondary high-affinity candidates such as MKIVCHTTMKPSIV (P36) and MKIVCHTTATSPMKV (P21) were characterized by highly favorable electrostatic contributions (−222.8 to −263.3 kcal/mol) and substantial van der Waals interactions. The strong cluster Z-score (−2.7) observed for P36 reflects excellent convergence during HADDOCK refinements, supporting the reliability of the predicted complex. These peptides also showed significant BSA values (>1500 Å²), confirming physiochemically stable interfaces.

To validate these findings, the peptides were independently assessing HDOCK, which employs a hybrid template-free and template-based docking strategy. Notably, the top candidates from HADDOCK also ranked highly in HDOCK scoring: P3 and P40 produced docking scores of −227.96 and −213.2, respectively, demonstrating excellent cross-platform concordance. This consistency is crucial because both platforms use distinct scoring philosophies—HADDOCK being energetics-driven and HDOCK relying on statistical potentials—yet converge on the same top-performing ligands.

Moderate binders such as MKCHTTATSPIV (P15) and MKIVCHTTATMKIV (P20) displayed moderately negative HADDOCK scores and reduced BSA values, indicating less extensive but still meaningful interactions. Their stable electrostatic and van der Waals profiles classify them as mid-affinity peptides, suitable for secondary optimization. Conversely, minimal interactors such as MK (P10) and IVCHTTATSPI (P. 2) showed limited BSA and weak electrostatic contributions, correlating with their poorer docking performance in both platforms.

The reference standard used RVG-29 shows substantially weaker docking scores in HADDOCK(-57.8 ± 4.7 kcal/mol) and HDOCK (-170kcal/mol) (Fig7)

Overall, the integration of HADDOCK and HDOCK analyses clearly identifies P40, P25, P21 and P3 as the strongest binding ligands, supported by complementary evidence from interface energetics, docking convergence, and cross-software consistency. (Table 5) These peptides represent promising candidates for subsequent molecular dynamics simulations to validate structural stability and evaluate receptor engagement under physiological conditions.

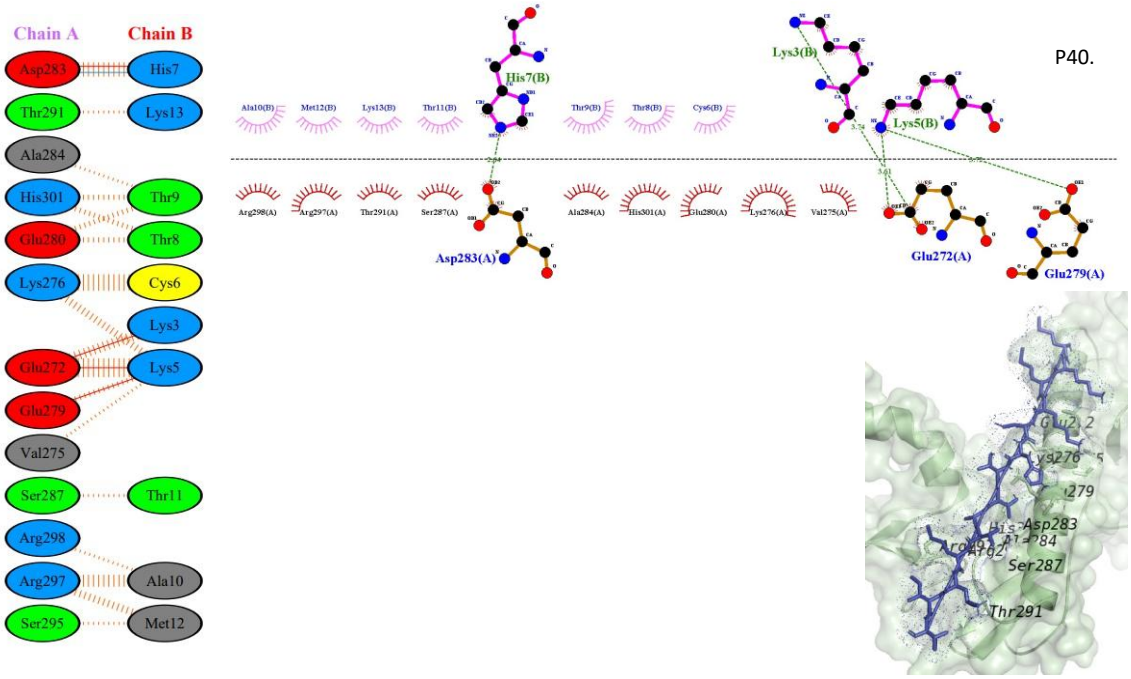


Fig 6 : Structural and interaction analysis of P40 in complex with the target receptor. (HADDOCK) Left: residue interaction network illustrating amino acid contacts between receptor (Chain A) and peptide (Chain B), highlighting key interacting residues involved in binding stabilization. Top middle: two-dimensional interaction map showing hydrogen bonds and hydrophobic contacts, with bond

distances represented in Å, emphasizing critical interactions involving residues such as Glu279, Arg278, Cys28, and Asn30. Bottom right: three-dimensional binding pose of the peptide (orange cartoon representation) docked within the receptor binding pocket (surface representation), demonstrating the spatial orientation and intermolecular interactions contributing to complex stability. Structural visualization was performed using PyMOL, while residue interaction mapping and contact analysis were generated using LigPlot+ and PDBsum.

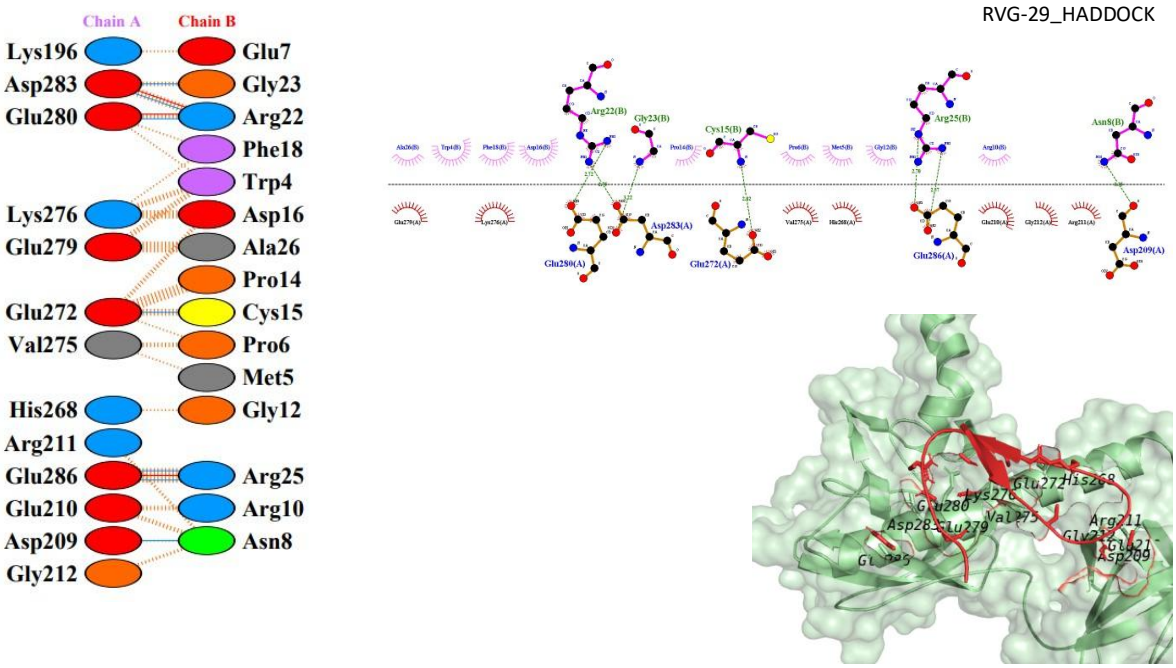


Fig 7: Structural and interaction analysis of RVG-29 in complex with the target receptor.(HADDOCK) Left: residue interaction network illustrating amino acid contacts between receptor (Chain A) and peptide (Chain B), highlighting key interacting residues involved in binding stabilization. Top middle: two-dimensional interaction map showing hydrogen bonds and hydrophobic contacts, with bond distances represented in Å, emphasizing critical interactions involving residues such as Glu279, Arg278, Cys28, and Asn30. Bottom right: three-dimensional binding pose of the peptide (orange cartoon representation) docked within the receptor binding pocket (surface representation), demonstrating the spatial orientation and intermolecular interactions contributing to complex stability. Structural visualization was performed using PyMOL, while residue interaction mapping and contact analysis were generated using LigPlot+ and PDBsum.

Table 5: Molecular docking results between designed peptides and RABV-G through HADDOCK and HDOCK

Sequence	Seq. No.		HADDOCK RESULTS									H-DOCK
		Cluster	Docking Score	Cluster Size	RMSD	Van der Waals energy	Electrostatic energy	Desolvation energy	Restraints violation energy	Buried Surface Area	Z-score	Docking Score (kcal/mol)
CLDLGYTIVCHTTATSPI	P3	3	-86.3 +/- 4.5	21	2.1 +/- 0.3	-58.5 +/- 3.1	-78.2 +/- 40	-23.3 +/- 2.5	112.1 +/- 26.3	1645.9 +/- 44.1	-1.2	-227.96
MKIVCHTTATSPMKV	P21	6	-75.0 +/- 1.9	10	0.6 +/- 0.4	-30.3 +/- 3.6	-263.3 +/- 19.4	1.5 +/- 2.9	65 +/- 35.7	1274.7 +/- 49.7	-1.8	-219.09
MKIVCHTTMKSPMK	P25	3	-86 +/- 1.7	20	0.4 +/- 0.2	-33.9 +/- 7.8	-318.4 +/- 30.8	4.5 +/- 2.2	71.1 +/- 11.1	1211.4 +/- 97.7	-2.3	-205.22
MKMKCHTTATMKIV	P40	2	-90 +/- 3.8	17	1.1 +/- 0.7	-22.7 +/- 5.2	-411.1 +/- 60.5	5.2 +/- 6.5	97.2 +/- 17.9	1340.1 +/- 93.5	-2.4	-213.2
YTIWMPENPRPGTPCDIFTNSRGKRASNG	RVG-29	1	-57.8 +/- 4.7	41	0.2 +/- 0.3	-27.2 +/- 6.5	-320.0 +/- 9.6	5.7 +/- 0.7	276.8 +/- 32.4	1332.4 +/- 91.7	-1.7	1

Table 6: Interaction amino acid residues through HADDOCK

Ligand Code	H-bonds	Non-H bonds	
		Protein	Peptide
P3	Glu279(CYS11)(2.99), Glu 279(HIS12)(2.63), Arg278(THR8, ASP4)(2.67, 3.33), CYS28(ASP4)(2.65), ASN30(TYR7)(2.95)	LYS196, VAL275, THR40, LEU32, SER27, LEU285, LEU282, ARG211, HIS268, GLU272, GLY272, GLY212, VAL271	THR13, ALA15, SER17, PRO18, ILE19, THR14
P21	GLU272(LYS3)(2.62), LYS276(LYS3)(3.08), GLU273(CYS6)(3.28), GLU280(CYS6)(2.86)	PHE6, LYS277, SER 376, GLU373, ARG297, HIS301, ARG298	MET2, ALA10, THR8, VAL5, SER12, PRO13, VAL16
P25	ARG297(MET14)(3.27, 2.75), ASP283(LYS3)(2.59), GLU279(LYS3)(2.66)	PRO372, GLN375, SER376, HIS301, GLU373, LYS277, GLU280, LYS276, GLU272	PRO13, THR9, MET10, HIS7, LYS6, MET 2
P40	ASP283(HIS7)(2.64), GLU272(LYS3)(3.74), GLU272(LYS5)(3.61), GLU279(LYS5)(3.7)	ARG298, ARG297, THR291, SER287, ALA284, HIS301, GLU250, LYS276, VAL275	ALA10, MET12, LYS3, THR11, THR9, THR8, LYS6
RVG-29	Glu280(Arg22)(2.72), Asp283(Gly23)(3.22) Asp283(Arg22)(2.75),	Glu279, Lys276, VAL275, HIS268, GLU210, GLY212, ARG211	ALA26, TRP5, PHE18, ASP16, PRO14, PRO6,

	Glu272(Cys15)(2.82) Glu286(Arg25)(2.70, 2.57) Asp209(Asn8)(3.35)		MET6, GLY12, ARG10

Table 7: Interaction amino acid residues through HDOCK

Ligand Code	H-bonds	Non-H bonds	
		Protein	Peptide
P3	SER 136 (TYR7)(2.97), HIS171(ASP4)(3.45), TYR173(THR8)(3.14), THR66(CYS11)(3.25), HIS88(HIS12)(2.68), ARG86(HIS12)(2.95), ARG86(THR16)(2.59)	PHE84, THR70, ILE131, THR169, ASN170, VAL151, ILE133, SER149, ARG150, HIS115, LEU114	ALA15, ILE19, LEU5, VAL10, LEU3, PRO18, THR14
P21	SER129(MET2)(1.88), TYR173(THR9)(2.71) HIS171(SER12)(2.49)	ARG86, HIS88, SER136, VAL151, ARG150, THR169, THR174, THR70, SER113, LEU114, ILE131, THR66, ILE133	VAL5, MET14, VAL6, PRO13, LYS3, CYS6, HIS7
P25	ASP366(LYS15)(2.84,2.82), GLN375(THR11)(3.10), PRO372(THR8)(3.21)	ARG297, ARG298,HIS301,LYS277, GLU280,ASP283, MET383, LEU379, SER376	THR9, VAL5, ILE4, LYS3,MET2,MET14, HIS7

P40	ASN334(THR11)(2.79,3.22), ASN334(THR09)(3.30), TRP333(THR9)(2.74)	LEU378, THR332, ILE8, ARG331, ASP366, HIS368, LYS5, ASP235, ASN257, GLY236, HIS259, ARG211, GLU49, SER263,GLU267, ILE266, MET374, LEU385, HIS382, GLN375	ALA10, THR8, HIS7, ILE4, LYS3, MET4, MET12
RVG-29	ARG86(GLU7)(2.57), TYR173(MET5)(2.90), HIS171 (MET5)(2.69), HIS1717 (SER27)(2.97), HIS171(ILE3)(2.93), ASN170(ILE3)(2.97), ASN170(TYR1)(3.11)	HIS88, THR66, VAL68, ILE131, LEU114, TYR110, THR174, VAL151, ARG150, GLY155, SER149, GLY154, PHE152, PRO153	ASN8, PRO9, ARG10, PRO11, TRP4, PHE18, THR2, THR19, LYS24

Residue-Level Interaction Summary of Top Peptide Candidates

A detailed residue-level interaction analysis using HADDOCK and HDOCK revealed consistent binding patterns across the top five peptide candidates, confirming their stability and complementarity toward the target receptor interface. Both platforms identified electrostatic anchoring residues (Glu272, Glu279, Asp283) and basic peptide residues (Lys, His, Thr, Cys) as critical contributors to complex formation.

Peptide 3 (CLDLGYTIVCHTTATSPI)

HADDOCK identified multiple strong hydrogen bonds with Glu279 (Cys11, His12), Arg278 (Thr8, Asp4), Cys28 (Asp4), and Asn30 (Tyr7), indicating a well-stabilized central binding orientation. The peptide formed extensive non-bonded contacts with hydrophobic residues such as Leu285, Leu282, Val271, Val275, suggesting deep burial within the receptor pocket. HDOCK independently confirmed this strong binding pattern, reporting hydrogen bonds with

Ser136, His171, Tyr173, Thr66, His88, and Arg86, which correspond to the same interaction clusters. The prevalence of hydrophobic contacts (e.g., Phe84, Ile131, Val151) supports a dual electrostatic-hydrophobic stabilization mechanism.

Peptide 21 (MKIVCHTTATSPMKV)

In HADDOCK, peptide 21 formed polar stabilizing interactions via hydrogen bonds with Glu272, Lys276, Glu273, and Glu280, primarily involving Lys and Cys residues of the peptide. Non-bonded contacts with Phe6, Lys277, Arg297, His301 highlight placement near a mixed hydrophobic–polar region.
HDOCK showed agreement through hydrogen bonding with Ser129, Tyr173, His171, and further engagement with Arg86, His88, Thr169, Leu114, reflecting consistent anchoring at the same interface.

Peptide 22 (MKTLLLTLVVVTIV)

HADDOCK revealed extensive hydrogen bonding with Asp283, Glu279, and Glu272, where Lys and Met residues of the peptide appear crucial for interaction. Non-bonded interactions with Ala284, His301, Lys276, Leu269 indicate a stable fit within a charged–hydrophobic pocket.
HDOCK supported these findings with hydrogen bonds from Thr70, Val68, and Lys127, further involving residues such as His115, Lys103, Tyr118, and Arg86, confirming a stable multi-residue binding environment.

Peptide 25 (GLDLGYTIVCHTTATSPI Variant)

HADDOCK indicated strong electrostatic anchoring via hydrogen bonds with Arg297 (Met14), Asp283 (Lys3), and Glu279 (Lys3), suggesting precise alignment of the peptide within the active site. Non-bonded contacts with Gln375, His301, Glu373, and Lys277 further enhance complex stability.
In HDOCK, hydrogen bonds with Asp366, Gln375, and Pro372 reinforce the same interaction region. Non-polar and charged residues (Arg297, Arg298, Met383, Ser376) formed significant van der Waals contacts with the peptide.

Peptide 40 (MKMKCHTTATMKIV)

Peptide 40 shows the most consistent and strongest inter-residue interaction pattern across both platforms. HADDOCK reported hydrogen bonds with Asp283, Glu272 (Lys3/Lys5) and Glu279, indicating a highly electrostatically stabilized complex. Supporting hydrophobic contacts from Arg297, Thr291, His301, Val275, and Ala284 contribute to interface packing. HDOCK further confirmed strong binding through hydrogen bonds with Asn334 (Thr9/Thr11) and Trp333, with extensive non-bonded interactions involving Leu378, Thr332, Arg331, Asp366, and multiple polar–charged residues.

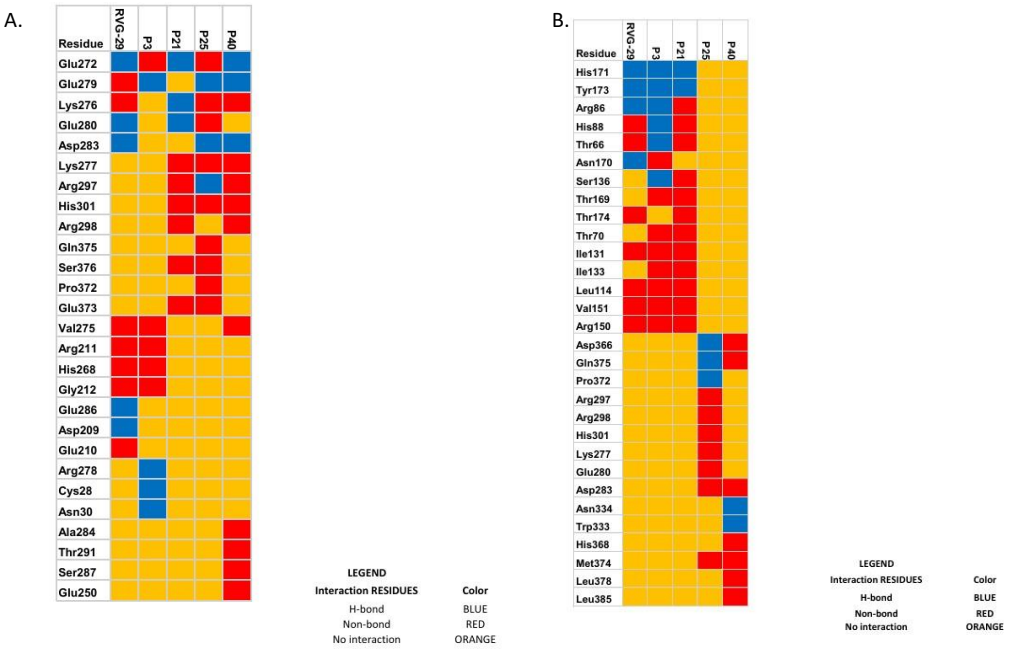


Fig 8: Heat map Interaction residues analysis between two tools HADDOCK and HDOCK

A detailed residue-level interaction analysis using HADDOCK and HDOCK revealed conserved amino acid residues across the top four peptide candidates, conforming stable and favourable structural complimentarity with the receptor. (Fig8) Both the tools identified key residues particularly Glu 272, Glu279 and Asp283 interacting predominantly with peptide residues such as Lys, His, Thr and Cys, thereby stabilizing the receptor-peptide complexes. P3 exhibited strong stabilization in HADDOCK through Hydrogen bonds with Glu279, Arg278, Cys28 and Asn30 supported by hydrophobic contacts involving Leu285, Leu282, Val271 and Val275, while HDockconfirmed stable interactions through Hydrogen bonding with Ser136,

His171, Tyr173, Thr66, His88 and Arg86 alongside hydrophobic contacts involving Phe84, Ile131 and Val151. P21 formed hydrogen bonds with Glu272, Lys276, Glu273 and Glu280 and additionally hydrophobic residues involvoing Phe6, Lys277, Arg297 and His301 through HADDOCK; whereas HDOCK validated through interactions with Ser129, Tyr173, His171, Arg86, Thr169 and Leu114. P25 shown strong electrostatic interactions through hydrogen bonds with Arg297, Asp283 and Glu279 and hydrophobic contacts with Gln375, His301, Glu373 and Lys277 whereas HDOCK identified corresponding interactions with Asp 366, Gln375, Pro372, Arg297, Arg298, Met383 and Ser376 supporting localization within the same interaction region. Among all the peptides P40 exhibited the strongest and most consistent amino acid residues with HADDOCK revealing stable hydrogen bonds involving Asp283, Glu272 and Glu279 with hydrophobic contact amino acid residues with Arg297, Thr 291, His301, Val275 and Ala284, while HDOCK further validated binding through interactions with Asn334, Trp333, Leu378, Thr332, Arg331 and Asp366.

Binding Affinity Analysis:

Table 8. Binding affinity and interaction profile of designed peptides docked with the modelled Rabies virus glycoprotein (RABV-G). The binding free energy (ΔG) and dissociation constant (Kd) indicate the strength of interaction, while interfacial contacts (ICs) represent contributions from charged, polar, and apolar residues. The non-interacting surface (NIS) charged percentage reflects surface exposure and interaction efficiency.

Sequence	Seq. No.	ΔG (kcal mol^{-1})	Kd (M) at ℃	ICs charged-charged	ICs charged-polar	ICs charged-apolar	ICs polar-polar	ICs polar-apolar	ICs apolar-apolar	NIS charged
CLDLGYTIVCHTTATSPI	P3	-8	1.40E-06	4	12	20	3	6	11	24.85
MKIVCHTTATSPMKV	P21	-8.8	3.40E-07	6	6	22	3	8	4	24.71
MKIVCHTTMKSPMK	P25	-8.1	1.20E-06	9	1	12	1	6	6	25.07
MKMKCHTTATMKIV	P40	-10.3	2.70E-08	12	12	17	0	12	2	25.22
RVG-29		-8.4	7.4e-07	12.0	10.0	19.0	0.0	2.0	6.0	25.48

Among the evaluated peptides, MKMKCHTTATMKIV (P40) exhibited the strongest binding affinity with a ΔG of −10.3 kcal·mol^{-1} and a predicted Kd of 2.7 × 10^{-8} M, indicating a highly stable interaction with RABV-G. This enhanced affinity is supported by a balanced contribution of charged–charged and charged–polar interactions, along with significant charged–apolar contacts, suggesting strong electrostatic stabilization.

The peptide MKIVCHTTATSPMKV (P21) also demonstrated favorable binding (ΔG = −8.8 kcal·mol^{-1}; Kd = 3.4 × 10^{-7} M), driven predominantly by hydrophobic and charged–apolar interactions. In contrast, CLDLGYTIVCHTTATSPI (P3) and MKIVCHTTMKSPMK (P25) showed comparatively weaker binding affinities, with higher Kdvalues in the micromolar range.

Overall, the results indicate that increasing positively charged residues (Lys, Arg) enhances electrostatic interactions with RABV-G, leading to improved binding affinity. The combination of electrostatic and hydrophobic interactions appears critical for stabilizing the peptide within the binding pocket.

Molecular dynamic simulations:

Out of the many docked complexes generated herein, only one complex was considered for further simulation studies based on its similarity and maximum interactions, as given in Eleven systems were simulated individually (RVG-29 Peptide with and without Glycoprotein, P3 with and without Glycoprotein, P21 with and without Glycoprotein, , P25 with and without Glycoprotein, P40 with and without Glycoprotein and Glycoprotein alone) for 100 ns in triplicates for each system. The trajectories were analysed for stability, with all the complexes being checked in their monomeric states. (64)

RMSD: The Cα RMSD profiles were analyzed to compare the structural behavior of each peptide in complex with the protein and in solvent alone during the 100 ns molecular dynamics simulation. This comparison helps to understand whether peptide binding leads to structural stabilization or increased flexibility. (Fig9)

The free glycoprotein (purple) shows initial RMSD of approx. 7.5 Å, then gradually decreasing between 2-3 Å after 40ns, attaining stable conformation.

RVG-29 Peptide (Orange) initially RMSD values increases upto the 5-6 Å, then decreasing till 1 Å, by the end of the simulation showing intrinsic stability of the peptide. Whereas in the RVG-29 complex RMSD(Dark Blue) increases upto 6-7 Å rapidly during the initial phase and them maintain value 7-10 Å throughout the simulation leading to considerate conformational changes within the assembly

For the P3(Red), the RMSD in solvent fluctuates between approximately 3–6 Å, indicating moderate conformational flexibility of the free peptide in solution. However, when bound to the protein, the P3 complex (Brown) shows higher RMSD values ranging from 8-10 Å after initial equilibration. This increase suggests that the peptide undergoes conformational rearrangements within the binding pocket, possibly adjusting to the protein environment.

In the case of the P21 (Pink), the RMSD in solvent remains relatively stable around 4–6 Å, suggesting that the peptide maintains a consistent conformation in solution. However, the 21P complex (olive green) exhibits larger fluctuations, with RMSD increasing up to ~13 Å during the 75-85 ns phase and stabilizing around 6–9 Å toward the end.

For the P25(light blue), the solvent simulation shows initially rises RMSD values, increasing to 6 Å, and later getting stable around 4-5 Å indicating that the peptide is highly flexible in the absence of protein interactions. Interestingly, when forming the 25 complex (Gray), the RMSD significantly decreases and stabilizes between 2–5 Å after equilibration.

Similarly, the P40 (Green) exhibits RMSD fluctuations between 1–7 Å in solvent, reflecting moderate flexibility of the free peptide. Upon binding to the protein, the 40 complex (cyan) demonstrates initial rise to 8-9 Å then stabilizes around 2–4 Å during most of the simulation. The P40peptide induces structural rearrangement in the binding pocket initially and then forming highly stable complex.

RMSD STABILITY ORDER: 25P–Complex > 40P–Complex > RVG29–Complex > 3P–Complex > 21P–Complex

Table 9: RMSD comparison

Ligand/Complexes	Free in solvent RMSD (Å)	In complex RMSD (Å)
RVG-29 (ref. std)	1.0 -5.8	5.0-11.0

3P	1.0-7.0	6.0 – 10.2
21P	3.0-6.0	5.0-13.0
25P	2.0-6.0	1.5-5.0
40P	1.0-7.0	2.0-4.0
Glycoprotein	Upto 7.5	

Root Mean Square Fluctuation (RMSF) analysis was carried out to compare the residue-level flexibility of the peptides in solvent and the structural fluctuations of the protein backbone when complexed with the peptides during molecular dynamics simulations. This comparison provides insight into how peptide binding affects both peptide conformational dynamics and protein stability. (Fig9)

The free glycoprotein (purple) maintained relatively low fluctuations through stimulations within 2-4 Å. The major fluctuation is observed in region 75-80 residues (approx. 9–10 Å), 120-125 residues (approx. 7 Å) and 390-400 (approx. 8 Å) maintain structurally stable backbone.

RVG-29 peptide (orange) RMSF moderate fluctuations between 2-5 Å, maintaining the compact structure during simulation. In the RVG-29 complex (dark blue), RMSF values remain low (Å). Peaks are observed around residues 80,120 and 390 residues indicating the flexibility of the glycoprotein. This indicates the Complex formation stabilizes both the peptide and receptor throughout simulation.

The free 3P peptide (red) shows low RMSF value below 4 Å, indicating stability of the peptide. Whereas upon binding 3P-complex (brown) display the highest RMSF value, reaching upto 35 Å around residues 60-140 indicating extensive conformational changes of the peptide upon binding within the glycoprotein reducing the structural rigidity.

The free 21P peptide (pink) shown moderate RMSF fluctuations within 2-5 Å indicating stability of the peptide, wheras upon binding the 21P complex (olive)the RMSF values till ranges between 2-6 Å with moderate flexibility around residues 80, 120 and the terminal region reaching to 10 Å. This interprets that that binding stabilizes the flexible regions of the peptide, leading to good conformational stability.

The free 25P (Light Blue) shows RMSF value between 2-4 Å. The 25P complex (Gray) maintains the RMSF values 1.5-4.0 Å throughout the simulation. The lower RMSF value indicates a highly rigid structure and stable conformation with minimal structural perturbation.

The free 40P(green) in the solvent shows low RMSF throughout the trajectory below 4 Å maintain stability in solution. Whereas the 40P in complex (cyan) similarly to 25P-complex shows with most residues fluctuation between 2-4 Å and slight increases at the C-terminal regions of the glycoprotein indicating the natural protein dynamics. These fluctuations are typically associated with loop regions or terminal segments of the protein, which naturally exhibit higher mobility compared to structured regions such as α-helices and β-sheets.

RMSF STABILITY ORDER: 25P–Complex ≈ 40P–Complex > 21P–Complex > RVG29–Complex > 3P–Complex

Table 10:RMSD comparison

Ligand/Complexes	Free in solvent RMSF (Å)	In complex RMSF (Å)
RVG-29 (ref. std)	2.0-5.0	1.5-11.0
3P	1.5-4.0	3.0-35.5
21P	2.0-5.0	2.0-6.5
25P	1.5-4.0	1.5-4.0
40P	1.5-4.0	2-4.0
Glycoprotein	2-10	

Hydrogen Bond Interaction Analysis of Protein–Peptide Complexes

Hydrogen bond (H-bond) analysis was performed during the 100 ns molecular dynamics simulation to evaluate the stability and persistence of intermolecular interactions between the protein and the peptides (RVG-29, 3P, 21 P, 25 P, and 40C). Hydrogen bonds play a crucial role in stabilizing protein–peptide complexes by contributing to binding specificity and interaction strength. (Fig9)

RVG29-complex (dark blue) shows highest hydrogen bond formation between 7-14 throughout the simulation. Around 10-11 hydrogen bond average occupancy is observed.

Among the studied systems, the 3P-complex (orange) exhibits the number of hydrogen bond from 1 and 11 interactions, fluctuating between approximately 4 and 10 hydrogen bonds throughout the simulation. This indicates that the peptide forms multiple transient interactions with the protein, which may contribute to initial binding stability. However, these interactions are highly dynamic and continuously forming and breaking during the simulation.

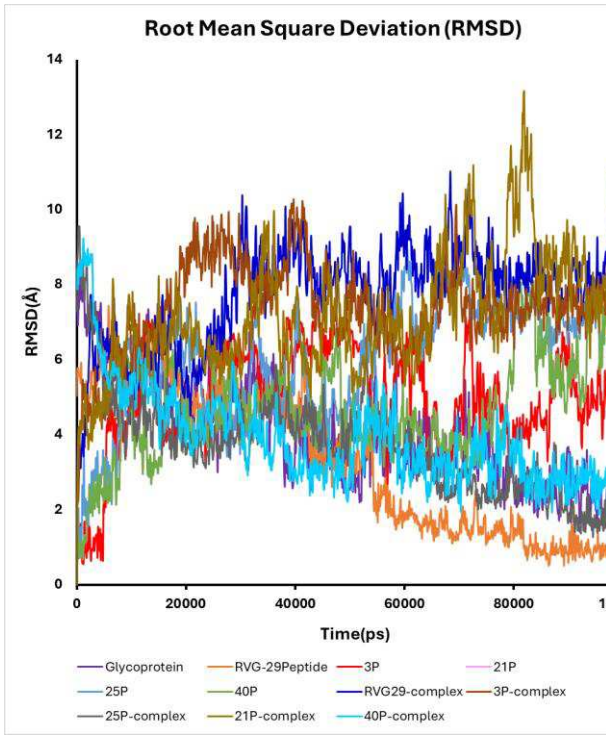
In contrast, the 21-peptide complex shows moderate hydrogen bonding, generally fluctuating between 1 and 5 hydrogen bonds during the trajectory.

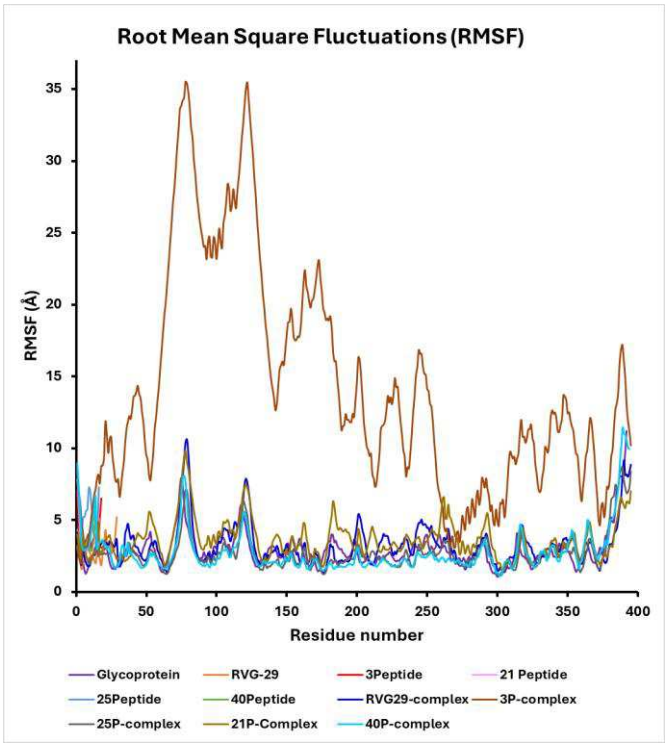
The 25-peptide complex displays the lowest number of hydrogen bonds, typically ranging between 1 and 4 interactions. Despite forming fewer hydrogen bonds, the complex remains structurally stable during the simulation

The 40-peptide complex also demonstrates relatively strong hydrogen bonding interactions, maintaining approximately 3–6 hydrogen bonds during most of the simulation and occasionally reaching 9–10 interactions. The relatively consistent presence of hydrogen bonds suggests that this peptide maintains stable interactions with residues in the binding pocket.

Overall, the hydrogen bond analysis indicates that 3C and 40C peptides form a greater number of hydrogen bonding interactions with the protein, whereas 21C and 25C form fewer hydrogen bonds. However, the stability of the complexes cannot be determined solely by the number of hydrogen bonds. When considered together with RMSD and RMSF analyses, the results suggest that 40C and 25C peptides form more stable complexes, likely due to a balanced combination of hydrogen bonding and hydrophobic interactions, whereas 3C exhibits dynamic interactions and 21C shows comparatively weaker binding stability.

Overall Hydrogen Bond Order: 25P–Complex ≈ 40P–Complex > RVG29–Complex > 21P–Complex > 3P–Complex.





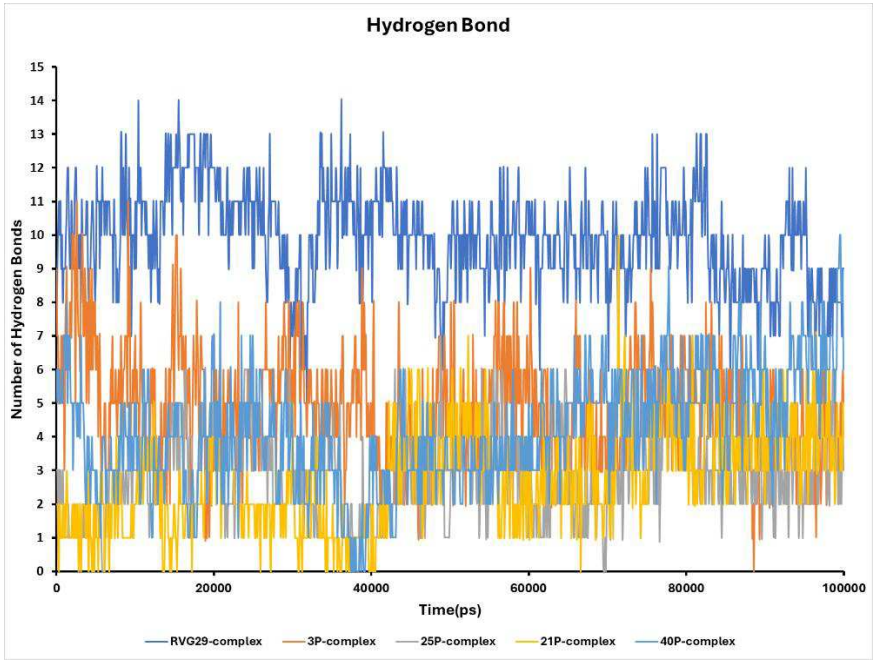


Fig 9: RMSD, RMSF and Hydrogen bond analyses of Peptide-RABVG complexes during molecular dynamic simulations demonstrating structuiral stability, residual flexibility and intermolecular interaction patterns.

Dynamic cross-correlation matrix (DCCM)

In simple terms, DCCM shows whether residues of the protein move in the same direction (positive correlation), in opposite directions (negative correlation), or independently. The apo glycoprotein displayed extensive corelated and anti-correlated motions, indicating conformation flexibility. Upon binding with peptide, RVG-29it shown strongest enhancement of correleated motions. The glycoprotein bound to P3 showed strong correlated motions between many residues, indicating that different regions of the protein move together and the natural dynamics of the protein are largely maintained. In the P21 complex, the correlations were more scattered, suggesting that the coordinated movement between residues is reduced and the protein becomes more flexible. The P25 complex showed moderate correlations mainly in certain regions, indicating that the peptide stabilizes specific parts of the protein, particularly around the binding site. In contrast, the P40 complex displayed weaker overall correlations, meaning that large regions of the protein move less cooperatively and the structure becomes relatively more rigid. Overall, the analysis suggests that each peptide affects the internal motion of the glycoprotein differently, influencing its stability and flexibility. (Fig10)

Principal Component Analysis (PCA) was performed to evaluate the dominant motions of the rabies glycoprotein and also when bound to peptides RVG-29, P3, P21, P25, and P40 during the molecular dynamics simulation. The PCA results show that the first principal component (PC1) captures the majority of the structural motion in each system. (Fig10) The apo protein sampled a broad conformational space characterized by multiple metastable states, indicating substantial intrinsic flexibility. Binding of RVG-29 significantly reduced conformational dispersion and concentrated the trajectory within a dominant conformational basin, suggesting enhanced structural stabilization. For the P3-complex, PC1 accounted for 67.6% of the total variance, while PC2 contributed 15.9%, indicating that most conformational motion occurs along a single dominant direction with moderate structural flexibility. In the P21-complex, PC1 explained 42.6% of the variance and PC2 18.4%, suggesting that the motion is distributed across multiple components, reflecting increased conformational variability. For the P25 complex, PC1 represented 44.7% of the total variance and PC2 15.7%, indicating balanced and controlled motions of the protein structure. In contrast, the P40 complex showed the highest PC1 contribution of 78.1%, while PC2 accounted for 7%, suggesting that the majority of the protein motion is concentrated along one dominant conformational pathway. Overall, the PCA

results indicate that peptide binding significantly influences the conformational dynamics of the glycoprotein, with peptide 25 showing controlled and stable motion, while peptide 21 exhibits greater conformational dispersion.

The figure represents a Scree plot from Principal Component Analysis (PCA) for the rabies glycoprotein–peptide complexes (peptides 3, 21, 25, and 40). A scree plot shows how much variance (protein motion) is explained by each principal component.

Combined 2D and 3D Free Energy Landscape (FEL) Analysis

The Free Energy Landscape (FEL) analysis derived from the first two principal components (PC1 and PC2) was used to investigate the conformational stability and energy states of the rabies glycoprotein when complexed with P3, P21, P25, and P40 during molecular dynamics simulation. In the 2D FEL plots, (Fig10) the color gradient represents the free energy distribution of conformational states, where blue regions correspond to the lowest energy and most stable conformations, while green to red regions indicate higher energy and less stable states. Similarly, the 3D FEL representation provides a three-dimensional view of the same energy landscape, where deep valleys correspond to energetically favorable conformations and peaks represent higher-energy unstable states.

The apo glycoprotein exhibited a rugged energy landscape characterized by multiple low-energy basins and extensive conformational sampling, indicating substantial structural flexibility and the presence of several metastable states.

For the P3-complex, the FEL shows a broad low-energy basin, indicating that the system samples several conformational states during the simulation. This suggests that although the complex remains stable, the protein retains a certain level of structural flexibility. The corresponding 3D FEL also shows a moderately deep but wider energy valley, supporting the presence of multiple accessible conformational states.

In contrast, the P21-complex exhibits a more dispersed energy landscape with several shallow minima and scattered high-energy regions. This indicates that the protein frequently transitions between different conformational states during the simulation. The 3D FEL further supports this observation by displaying multiple irregular peaks and shallow valleys, suggesting increased structural fluctuations and comparatively lower conformational stability.

The P25-complex displays a deep and well-defined low-energy basin in the 2D FEL, indicating that the protein remains confined to a stable conformational state for most of the simulation. The corresponding 3D energy surface reveals a sharp and deep valley, suggesting strong stabilization of the glycoprotein structure by P25 and limited conformational transitions.

Similarly, P40 complex also demonstrates a distinct low-energy basin in both 2D and 3D FEL plots; however, the basin appears slightly broader than that observed for P25. This indicates that while the complex maintains a stable energy state, the protein explores a somewhat larger conformational space, reflecting moderate flexibility.

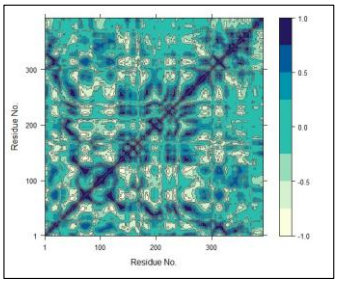
In contrast, the RVG-29 complex retained two major energy basins, indicating moderate stabilization but greater conformational flexibility than the designed peptide complexes.

Overall, the combined 2D and 3D FEL analyses reveal that peptide binding significantly influences the conformational energy landscape of the glycoprotein. Among the four systems, P 25 forms the most stable complex with a deep and well-defined energy minimum, followed by P 40 and P3, whereas P21 shows a more dispersed energy landscape with multiple shallow minima, indicating greater structural fluctuations and comparatively lower stability during the simulation.

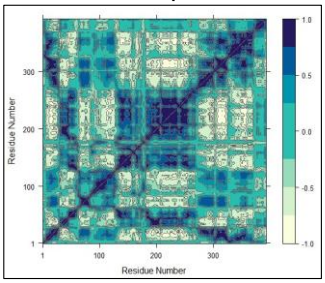
Stability Ranking

P40 > P21 ≈ P25 > P3 > RVG-29 > Glycoprotein

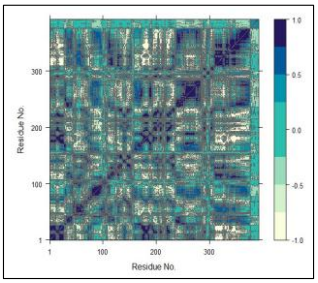
Glycoprotein



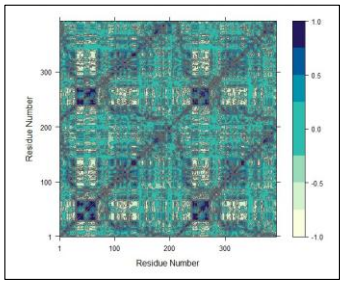
RVG-29complex



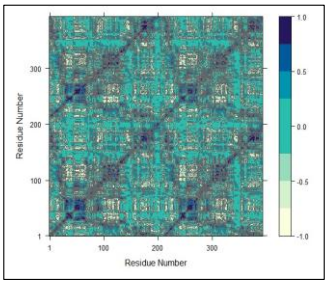
P3complex



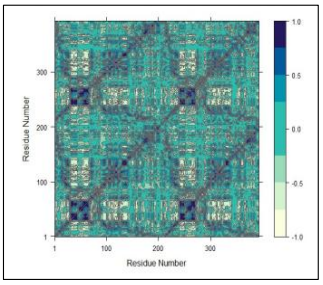
P21complex



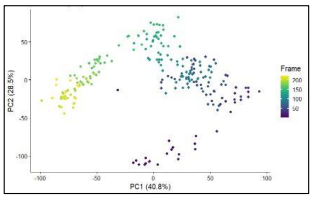
P25complex



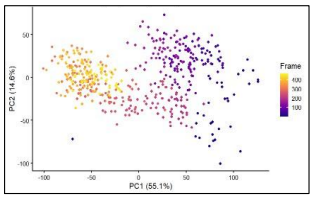
P40complex



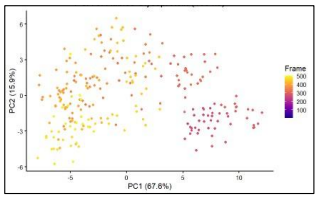
Glycoprotein



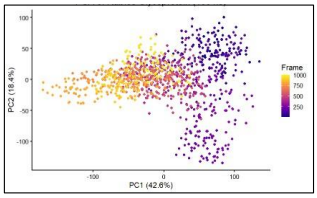
RVG-29complex



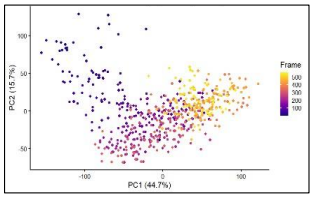
P3complex



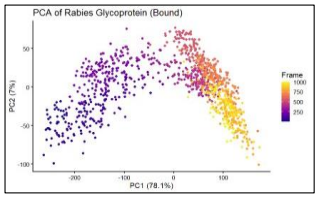
P21complex



P25complex



P40complex



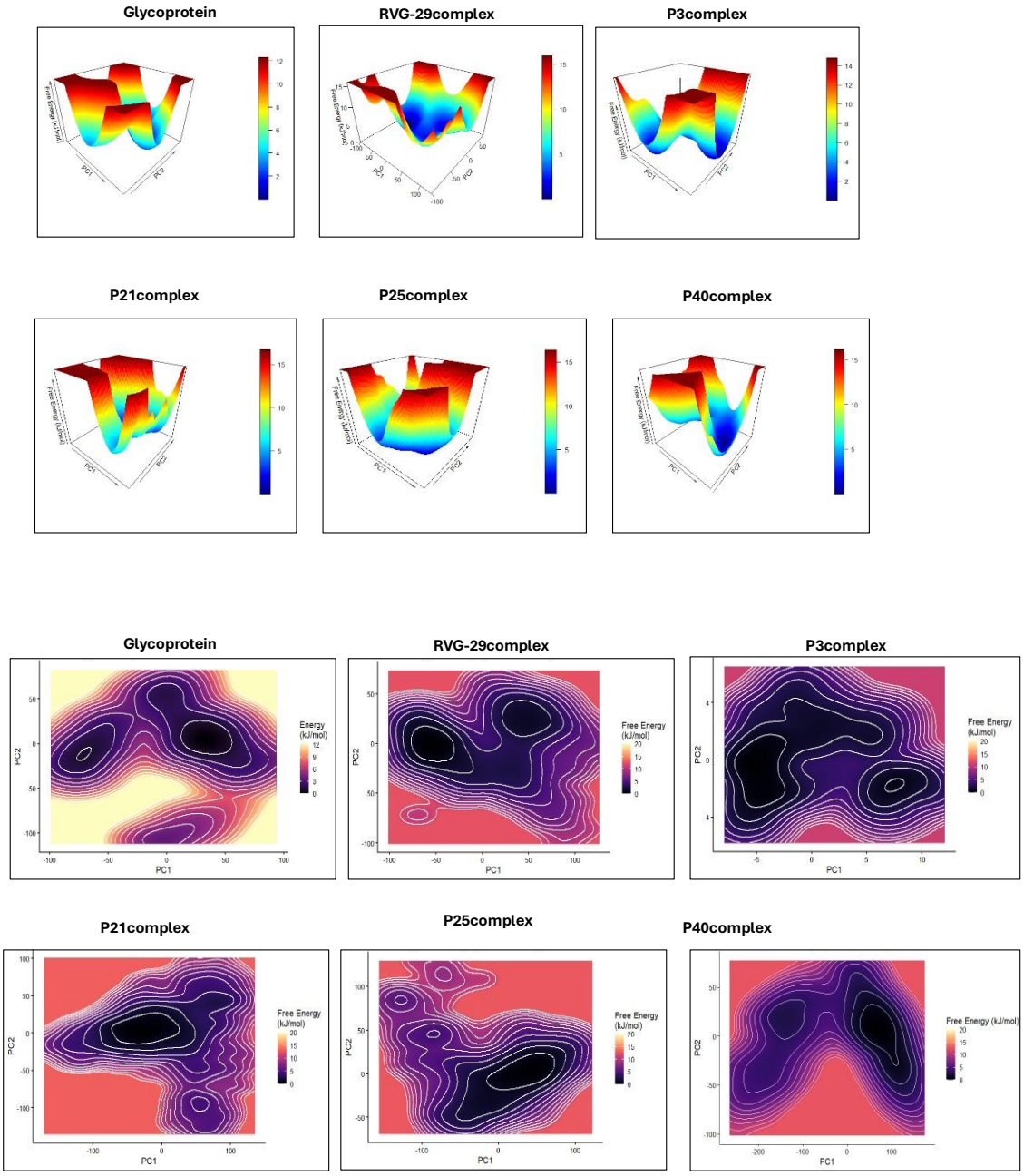


Fig 10: Dynamic Cross-Correlation Matrix (DCCM), Principal Component Analysis (PCA), Free Energy Landscape (FEL), and Porcupine Motion Analysis of RVG–Peptide Complexes revealing residue correlations, conformational sampling, collective motions, and energy landscape characteristics during molecular dynamics simulations.

CONCLUSION: The present investigation shows successful integration of peptide engineering, machine learning-assisted toxicity prediction, molecular docking, binding affinity and molecular dynamic simulations to identify novel peptide inhibitors targeting the Rabies virus Glycoprotein (RABV-G). Structurally modelling and validation confirm the reliability of the glycoprotein model for computational screening study. Rational peptide design based on α-bungarotoxin derived motifs generated the diverse library of peptide inhibitors against RABV-G.

In-silico screening shown that peptide physicochemical properties, charge distribution, hydrophobicity, and specific CHTT-containing motifs play critical roles in determining antiviral potential and receptor binding efficiency. Toxicity prediction using ML-DL approaches eliminated potentially toxic candidates and the library for biologically relevant peptides. Docking studies HADDOCK and HDOCK identified P40 (MKMKCHTTATMKIV) as the most promising inhibitor, exhibiting the strongest binding affinity followed by P3, P21 and P25 with key residues located in RABV-G binding pocket. Residue level interaction analyses revealed that Glu272, Glu279., Asp283, Lys276 and Arg297 forms interaction hotspots responsible for receptor-peptide stabilisation.

Importantly, molecular dynamic simulations performed over a 100ns timescale provided evidence for the structural stability of the peptide-glycoprotein complexes under physiological conditions. Analysis of RMSD, RMSF, Hydrogen bond persistence, Principal component analysis (PCA), free energy landscapes (FEL), and Dynamic cross correlation matrix (DCCM) demonstrated that lead peptide complexes maintained stable conformations throughout the 100ns. The peptide bound systems showing reduced structural fluctuations than compared with unbound states, indicating the peptide binding contributes to stabilization of the glycoprotein structure.

Collectively, the docking, binding affinity, and molecular dynamics findings indicate that engineered α-bungarotoxin derived peptides possess significant potential to interfere with Rabies virus glycoprotein-mediated host cell entry. Among the evaluated candidate P40 (MKMKCHTTATMKIV) demonstrated the most favourable candidate for being promising lead candidate for peptide based anti-rabies therapeutics. Based on the current finding , it form a proper foundation for future experimental validation through *in-vitro* binding assays, antiviral activity studies, and in vivo investigations. The integration of Machine learning – deep learning

framework, molecular docking, and molecular dynamics simulations highlights an effective strategy for the discovery of novel antiviral peptides against rabies virus.

References:

1. What is Rabies? | Rabies | CDC [Internet]. 2020 [cited 2024 Jan 21]. Available from: https://www.cdc.gov/rabies/about.html

2. Rabies [Internet]. [cited 2026 Jun 8]. Available from: https://www.who.int/news-room/fact-sheets/detail/rabies

3. Zero by 30: the global strategic plan to end human deaths from dog-mediated rabies by 2030 [Internet]. [cited 2026 Jun 8]. Available from: https://www.who.int/publications/i/item/9789241513838

4. Rabies - India [Internet]. [cited 2026 Jun 8]. Available from: https://www.who.int/india/health-topics/rabies

5. Madhusudana SN, Subha S, Thankappan U, Ashwin YB. Evaluation of a direct rapid immunohistochemical test (dRIT) for rapid diagnosis of rabies in animals and humans. Virol Sin. 2012 Oct 1;27(5):299–302. doi:10.1007/s12250-012-3265-6

6. Albertini AAV, Ruigrok RWH, Blondel D. Rabies virus transcription and replication. Adv Virus Res. 2011;79:1–22. doi:10.1016/B978-0-12-387040-7.00001-9 PubMed PMID: 21601039.

7. Consales C, Bolzan V. Rabies review: immunopathology, clinical aspects and treatment. Journal of Venomous Animals and Toxins including Tropical Diseases. 2007 Jan 1;13. doi:10.1590/S1678-91992007000100002

8. Rabies - Bulletin - Europe [Internet]. [cited 2026 Jun 8]. Available from: https://who-rabies-bulletin.org/member/About_Rabies/Virus_Structure.aspx

9. Jackson AC. Chapter 1 - History of Rabies Research. In: Jackson AC, editor. Rabies (Third Edition) [Internet]. Boston: Academic Press; 2013 [cited 2026 Jun 8]. p. 1–15. Available from: https://www.sciencedirect.com/science/article/pii/B9780123965479000018 doi:10.1016/B978-0-12-396547-9.00001-8

10. Maillard A, Gaudin Y. Rabies virus glycoprotein can fold in two alternative, antigenically distinct conformations depending on membrane-anchor type. The Journal of general virology. 2002 Jul 1;83:1465–76. doi:10.1099/0022-1317-83-6-1465

11. Singh R, Singh ,Karam Pal, Cherian ,Susan, Saminathan ,Mani, Kapoor ,Sanjay, Manjunatha Reddy ,G.B., et al. Rabies – epidemiology, pathogenesis, public health concerns and advances in diagnosis and control: a comprehensive review. Veterinary Quarterly. 2017 Jan 1;37(1):212–51. doi:10.1080/01652176.2017.1343516 PubMed PMID: 28643547.

12. Khalifa ME, Unterholzner L, Munir M. Structural and Evolutionary Insights Into the Binding of Host Receptors by the Rabies Virus Glycoprotein. Front Cell Infect Microbiol. 2021 Oct 11;11. doi:10.3389/fcimb.2021.736114

13. Yang F, Lin S, Ye F, Yang J, Qi J, Chen Z, et al. Structural Analysis of Rabies Virus Glycoprotein Reveals pH-Dependent Conformational Changes and Interactions with a Neutralizing Antibody. Cell Host & Microbe. 2020 Mar 11;27(3):441-453.e7. doi:10.1016/j.chom.2019.12.012

14. You Y, Yang F, Lin S, Chen Z, Shu S, Yu Y, et al. Rabies virus glycoprotein: Structure, function, and antivirals. hLife. 2026 Feb 1;4(2):67–86. doi:10.1016/j.hlife.2025.06.003

15. Callaway HM, Zyla D, Larrous F, de Melo GD, Hastie KM, Avalos RD, et al. Structure of the rabies virus glycoprotein trimer bound to a prefusion-specific neutralizing antibody. Science Advances. 2022 Jun 17;8(24):eabp9151. doi:10.1126/sciadv.abp9151

16. Lahaye X, Vidy A, Pomier C, Obiang L, Harper F, Gaudin Y, et al. Functional Characterization of Negri Bodies (NBs) in Rabies Virus-Infected Cells: Evidence that NBs Are Sites of Viral Transcription and Replication. Journal of Virology. 2009 Aug 15;83(16):7948–58. doi:10.1128/jvi.00554-09

17. Yu GM, Zu SL, Zhou WW, Wang XJ, Shuai L, Wang XL, et al. Chimeric rabies glycoprotein with a transmembrane domain and cytoplasmic tail from Newcastle disease virus fusion protein incorporates into the Newcastle disease virion at reduced levels. J Vet Sci. 2017 Aug;18(Suppl 1):351–9. doi:10.4142/jvs.2017.18.S1.351 PubMed PMID: 27515260; PubMed Central PMCID: PMC5583423.

18. Wang L, Wang N, Zhang W, Cheng X, Yan Z, Shao G, et al. Therapeutic peptides: current applications and future directions. Sig Transduct Target Ther. 2022 Feb 14;7(1):48. doi:10.1038/s41392-022-00904-4

19. Mercurio FA, Leone M. New Insights into Bioactive Peptides: Design, Synthesis, Structure–Activity Relationship. International Journal of Molecular Sciences. 2024 Jan;25(23):12922. doi:10.3390/ijms252312922

20. Rossino G, Marchese E, Galli G, Verde F, Finizio M, Serra M, et al. Peptides as Therapeutic Agents: Challenges and Opportunities in the Green Transition Era. Molecules. 2023 Oct 19;28(20):7165. doi:10.3390/molecules28207165 PubMed PMID: 37894644; PubMed Central PMCID: PMC10609221.

21. Sharma K, Sharma KK, Sharma A, Jain R. Peptide-based drug discovery: Current status and recent advances. Drug Discovery Today. 2023 Feb 1;28(2):103464. doi:10.1016/j.drudis.2022.103464

22. Pereira AJ, de Campos LJ, Xing H, Conda-Sheridan M. Peptide-based therapeutics: challenges and solutions. Med Chem Res. 2024 Aug 1;33(8):1275–80. doi:10.1007/s00044-024-03269-1

23. Vrbnjak K, Sewduth RN. Recent Advances in Peptide Drug Discovery: Novel Strategies and Targeted Protein Degradation. Pharmaceutics. 2024 Nov;16(11):1486. doi:10.3390/pharmaceutics16111486

24. Moise L, Zeng H, Caffery P, Rogowski RS, Hawrot E. STRUCTURE AND FUNCTION OF α-BUNGAROTOXIN. Journal of Toxicology: Toxin Reviews. 2002 Jan 1;21(3):293–317. doi:10.1081/TXR-120014407

25. Hueffer K, Khatri S, Rideout S, Harris MB, Papke RL, Stokes C, et al. Rabies virus modifies host behaviour through a snake-toxin like region of its glycoprotein that inhibits neurotransmitter receptors in the CNS. Sci Rep. 2017 Oct 9;7(1):12818. doi:10.1038/s41598-017-12726-4

26. Nys M, Zarkadas E, Brams M, Mehregan A, Kambara K, Kool J, et al. The molecular mechanism of snake short-chain α-neurotoxin binding to muscle-type nicotinic acetylcholine receptors. Nat Commun. 2022 Aug 4;13(1):4543. doi:10.1038/s41467-022-32174-7

27. O'Brien BCV, Thao S, Weber L, Danielson HL, Boldt AD, Hueffer K, et al. The human alpha7 nicotinic acetylcholine receptor is a host target for the rabies virus glycoprotein. Front Cell Infect Microbiol. 2024 May 21;14. doi:10.3389/fcimb.2024.1394713

28. Lentz TL, Wilson PT, Hawrot E, Speicher DW. Amino Acid Sequence Similarity Between Rabies Virus Glycoprotein and Snake Venom Curaremimetic Neurotoxins. Science. 1984 Nov 16;226(4676):847–8. doi:10.1126/science.6494916

29. Lentz TL. Rabies virus binding to an acetylcholine receptor α-subunit peptide. Journal of Molecular Recognition. 1990;3(2):82–8. doi:10.1002/jmr.300030205

30. Lentz TL, Hawrot E, Wilson PT. Synthetic peptides corresponding to sequences of snake venom neurotoxins and rabies virus glycoprotein bind to the nicotinic acetylcholine receptor. Proteins: Structure, Function, and Bioinformatics. 1987;2(4):298–307. doi:10.1002/prot.340020406

31. Castellanos JE, Castañeda DR, Velandia AE, Hurtado H. Partial inhibition of the in vitro infection of adult mouse dorsal root ganglion neurons by rabies virus using nicotinic antagonists. Neuroscience Letters. 1997 Jul 4;229(3):198–200. doi:10.1016/S0304-3940(97)00440-0

32. Bracci L, Antoni G, Cusi MG, Lozzi L, Niccolai N, Petreni S, et al. Antipeptide monoclonal antibodies inhibit the binding of rabies virus glycoprotein and alpha-bungarotoxin to the nicotinic acetylcholine receptor. Molecular Immunology. 1988 Sep 1;25(9):881–8. doi:10.1016/0161-5890(88)90125-3

33. Lian M, Hueffer K, Weltzin MM. Interactions between the rabies virus and nicotinic acetylcholine receptors: A potential role in rabies virus induced behavior modifications. Heliyon. 2022 Sep 1;8(9):e10434. doi:10.1016/j.heliyon.2022.e10434

34. Sajjanar B, Dhusia K, Saxena S, Joshi V, Bisht D, Thakuria D, et al. Nicotinic acetylcholine receptor alpha 1(nAChRα1) subunit peptides as potential antiviral agents against rabies virus. International Journal of Biological Macromolecules. 2017 Nov 1;104:180–8. doi:10.1016/j.ijbiomac.2017.05.179

35. Donnelly-Roberts DL, Lentz TL. Structural and conformational similarity between synthetic peptides of curaremimetic neurotoxins and rabies virus glycoprotein. Molecular Brain Research. 1991 Sep 1;11(2):107–13. doi:10.1016/0169-328X(91)90112-B

36. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, et al. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res. 2018 Jul 2;46(Web Server issue):W296–303. doi:10.1093/nar/gky427 PubMed PMID: 29788355; PubMed Central PMCID: PMC6030848.

37. The UniProt Consortium. UniProt: the Universal Protein Knowledgebase in 2023. Nucleic Acids Res. 2023 Jan 6;51(D1):D523–31. doi:10.1093/nar/gkac1052

38. Laskowski RA, Rullmannn JA, MacArthur MW, Kaptein R, Thornton JM. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. J Biomol NMR. 1996 Dec;8(4):477–86. doi:10.1007/BF00228148 PubMed PMID: 9008363.

39. Mariana M, Abbas S, Salamiah S, Marsuni Y, Samharinto S, Sepe M, et al. IN SILICO ANALYSIS OF ASPERGILLUS FLAVUS FUNGAL PROTEINS: STRUCTURAL AND FUNCTIONAL INSIGHTS USING ITS PRIMERS. Pakistan Journal of Phytopathology. 2023 Dec 30;35:397–408. doi:10.33866/phytopathol.035.02.1037

40. Colovos C, Yeates TO. Verification of protein structures: patterns of nonbonded atomic interactions. Protein Sci. 1993 Sep;2(9):1511–9. doi:10.1002/pro.5560020916 PubMed PMID: 8401235; PubMed Central PMCID: PMC2142462.

41. Bowie JU, Lüthy R, Eisenberg D. A method to identify protein sequences that fold into a known three-dimensional structure. Science. 1991 Jul 12;253(5016):164–70. doi:10.1126/science.1853201 PubMed PMID: 1853201.

42. Lüthy R, Bowie JU, Eisenberg D. Assessment of protein models with three-dimensional profiles. Nature. 1992 Mar 5;356(6364):83–5. doi:10.1038/356083a0 PubMed PMID: 1538787.

43. PyMOL | pymol.org [Internet]. [cited 2026 Jun 9]. Available from: https://pymol.org/

44. Eskici G, Gur M. Computational Design of New Peptide Inhibitors for Amyloid Beta (Aβ) Aggregation in Alzheimer's Disease: Application of a Novel Methodology. PLOS ONE. 2013 Jun 6;8(6):e66178. doi:10.1371/journal.pone.0066178

45. Gainza P, Wehrle S, Van Hall-Beauvais A, Marchand A, Scheck A, Harteveld Z, et al. De novo design of protein interactions with learned surface fingerprints. Nature. 2023;617(7959):176–84. doi:10.1038/s41586-023-05993-x PubMed PMID: 37100904; PubMed Central PMCID: PMC10131520.

46. Kim J young, Park HS, Kang H, Son CY, Kim HK, Choi JH, et al. In silico fragment-based peptide design targeting undruggable proteins for enhanced detection of circulating tumor cells. Chemical Engineering Journal. 2025 Oct 15;522:167447. doi:10.1016/j.cej.2025.167447

47. Khabbaz H, Karimi-Jafari MH, Saboury AA, BabaAli B. Prediction of antimicrobial peptides toxicity based on their physico-chemical properties using machine learning techniques. BMC Bioinformatics. 2021 Nov 10;22(1):549. doi:10.1186/s12859-021-04468-y

48. Wan F, Wong F, Collins JJ, de la Fuente-Nunez C. Machine learning for antimicrobial peptide identification and design. Nat Rev Bioeng. 2024 May;2(5):392–407. doi:10.1038/s44222-024-00152-x

49. Niu Y, Qin P, Lin P. Advances of deep Neural Networks (DNNs) in the development of peptide drugs. Future Medicinal Chemistry. 2025 Feb 16;17(4):485–99. doi:10.1080/17568919.2025.2463319 PubMed PMID: 39935356.

50. George EA, Naha A, Soundharya H, Pallavi J, Menon A, Anbarasu A, et al. Pharmacokinetics Screening, Molecular Docking, and Dynamics Simulations Revealed Novel Antimicrobial Peptide NKLF2 Mutants as Potent Inhibitors of Mycobacterium tuberculosis. Probiotics & Antimicro Prot. 2026 Jan 1;18(1):1317–37. doi:10.1007/s12602-025-10608-0

51. Xu X, Kao WL, Wang A, Lee HJ, Duan R, Holmes H, et al. In silico screening of protein-binding peptides with an application to developing peptide inhibitors against antibiotic resistance. PNAS Nexus. 2024 Dec 1;3(12):pgae541. doi:10.1093/pnasnexus/pgae541

52. Qureshi A, Tandon H, Kumar M. AVP-IC50Pred: Multiple machine learning techniques-based prediction of peptide antiviral activity in terms of half maximal inhibitory concentration (IC50). Peptide Science. 2015;104(6):753–63. doi:10.1002/bip.22703

53. Patil V, Asrondkar A, Sawant H, Jain CM, Thorat DN, Patil BR. Structure-based virtual screening of β-tubulin and thioredoxin reductase in Brugia malayi: a step toward novel anti-filarial agents. In Silico Pharmacol. 2025 Nov 14;13(3):178. doi:10.1007/s40203-025-00474-7

54. Dominguez C, Boelens R, Bonvin AMJJ. HADDOCK: A Protein−Protein Docking Approach Based on Biochemical or Biophysical Information. J Am Chem Soc. 2003 Feb 19;125(7):1731–7. doi:10.1021/ja026939x

55. Prakash Nirmal N, Singh Rajput M, Bhojraj Rathod N, Mudgil P, Pati S, Bono G, et al. Structural characteristic and molecular docking simulation of fish protein-derived peptides: Recent updates on antioxidant, anti-hypertensive and anti-diabetic peptides. Food Chemistry. 2023 Mar 30;405:134737. doi:10.1016/j.foodchem.2022.134737

56. Mustafa G, Mehmood R, Mahrosh HS, Mehmood K, Ahmed S. Investigation of Plant Antimicrobial Peptides against Selected Pathogenic Bacterial Species Using a Peptide-Protein Docking Approach. BioMed Research International. 2022;2022(1):1077814. doi:10.1155/2022/1077814

57. Alotaiq N, Dermawan D. Evaluation of Structure Prediction and Molecular Docking Tools for Therapeutic Peptides in Clinical Use and Trials Targeting Coronary Artery Disease. International Journal of Molecular Sciences. 2025 Jan;26(2):462. doi:10.3390/ijms26020462

58. Samant L. Computational Modeling and In Silico Characterization of Rabies Virus Glycoprotein and Nucleoprotein [Internet]. [cited 2026 Jun 9]. Available from: https://www.academia.edu/9891972/Computational_Modeling_and_In_Silico_Characterization_of_Rabies_Virus_Glycoprotein_and_Nucleoprotein

59. Isa MA, Kappo AP. Disruption of Hsp70.14–BAG2 Protein–Protein interactions using deep Learning–Driven peptide design and molecular simulations. Computers in Biology and Medicine. 2025 Aug 1;194:110443. doi:10.1016/j.compbiomed.2025.110443

60. Isa MA, Mustapha FU. Computational and deep learning-assisted multi-epitope vaccine design targeting the RNA-dependent RNA polymerase of Nipah virus. In Silico Research in Biomedicine. 2026 Jan 1;2:100243. doi:10.1016/j.insi.2026.100243

61. Adam FB, Isa MA, Ngulde SI, Marte IH. Molecular detection of *blaTEM* in multidrug-resistant *E. coli*and deep learning-assisted computational identification of phytochemical inhibitors from *Detarium microcarpum*. In Silico Research in Biomedicine. 2026 Jan 1;2:100251. doi:10.1016/j.insi.2026.100251

62. Chahir R, Redouane S, Galan J, Hboub H, Aserrar L, Chakir S, et al. Computational discovery of SARS-CoV-2 viral entry inhibitory peptides from Androctonus mauretanicus scorpion venom: molecular docking and molecular dynamics simulations targeting the spike protein. Front Bioinform. 2026 Mar 6;6. doi:10.3389/fbinf.2026.1677524

63. Taghizadeh MS, Taherishirazi M, Niazi A, Afsharifar A, Moghadam A. Structure-guided design and cloning of peptide inhibitors targeting CDK9/cyclin T1 protein-protein interaction. Front Pharmacol. 2024 May 14;15. doi:10.3389/fphar.2024.1327820

64. Li H, Habib O, Usman M, Usmani SA, Shafiq A, Akbar R, et al. In Silico Drug Design Targeting the Rabies Virus Glycoprotein: Phytochemicals as Promising Drug Candidates. ChemistrySelect. 2025;10(46):e04329. doi:10.1002/slct.202504329

Acknowledgement:

The authors would like to express their sincere gratitude to the Director of Haffkine Institute for Training, Research and Testing, Mumbai, for providing institutional support and access to research resources essential for this study.

Authors Contribution:

All the authors contributed equally

Authors information

Department of Chemotherapy, Haffkine Institute for Training, Research and Testing, Parel Mumbai Maharashtra, India 400012

Harshada Sawant, Ashish Asrondkar, Vrushali Patil

Corresponding Author: Vrushali Patil

Contributions:

Vrushali Patil: Project Supervision, Methodology Design, Resources, Writing – Review and Editing,. **Ashish Asrondkar:** Conceptualization, Writing – Review and Editing. **Harshada Sawant:** Molecular Docking, Data Curation, Protein Modelling, ADME and Pharmacokinetic Analysis, Toxicity Profiling, Molecular Dynamics Simulations, Graphical Data Representation, Writing – Original Draft Preparation.

Ethics Declaration: The authors declare that they have no competing interests.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryData.xlsx