In Silico Identification of Novel Ligand Molecules for Rabies Nucleoprotein using Structure-Based Method

Muthukumar Balamurugan*, Aril Bhoomigowda, Tanvee Pardesdhi, Sameer Choudhary, Shama Mujawer

Research and Development, Rasa Life Science Informatics, Pune, Maharastra-411016, INDIA.

ABSTRACT

Background: Rabies is a life-threatening condition which much privileged in Asia and Africa causes high mortality annually. It has to be treated before showing of any clinical signs. Presently vaccination is the lone accessible medication. Several research are taking place around worldwide for production of epitope based and DNA vaccines. Rabies Nucleoprotein plays a decisive role in transcription and replication process in rabies virus. In this study 2 ligand molecules were identified to prevent transcription and replication of rabies virus. **Methods:** Three dimensional structure of nucleoprotein model was modeled using Phyre2 server. The protein structure validation was carried out in Ramachandran plot by RAMPAGE. Energy minimization was performed for molecular dynamics study which was proceeded using Gromacs 5.1. Structure based pharmacophore method was employed to identify the ligand molecules against nucleoprotein by Ligand Scout. Molecular docking studies was executed in Auto Dock tools. Loop docking scheme has done for validation of binding energy. **Results:** 12 molecules were presenting reasonable binding energy interaction with nucleoprotein. Two ligand compounds ZINC10530604 and ZINC10530605 were exhibit virtuous binding energy -8.0 and -7.3 respectively. **Conclusion:** The identified ligands can be potential drug compound against rabies nucleoprotein by either preventing or inhibiting of transcriptional and translational activity.

Key words: Nucleoprotein, Homology modeling, Energy minimization and Simulation, Structure-based pharmacophore, Molecular docking.

BACKGROUND

Rabies, a fatal encephalitis neurotropic disease caused by Rabies Virus (RV) in mammals including humans. RV, a single stranded negative sense RNA virus belongs to mononegavirales order, lyssaviridae genus of rhabdoviridae family.[1] Without proper vaccination it causes more than 55000 deaths annually in Asia and Africa.[2] A bullet shaped virion composed of five proteins namely Glycoprotein (G), the transmembrane protein involves in interaction with host cell receptor.[3] Nucleoprotein (N), a RNA encapsidating protein by binding with Phosphoprotein (P), in multifunctional role, function as chaperone for forming N-P complex to specifically encapsidating to genomic RNA and also a cofactor molecule for large protein (L) represent RNA dependent RNA polymerase (RdRp).[4] RV N protein encapsulate viral genome with the help of P and L protein forming ribonucleoprotein (RNP) complex.[1] Matrix protein (M) plays a role in viral budding.[5] Nucleoprotein plays a decisive role in transcription and replication in rhabdoviridae family of viruses.[6,8-9] Vesicular stomatitis virus (VSV) another notable virus belonging to rhabdoviridae family showed highly conserved structural and functional characteristic features with RV. Even though there is lacking sequence homology between RV N and VSV N, they are showing highly conserved characters of protein folding, RNA binding and assembly. The structural properties of nucleoprotein revealed that RV N is made up of 450 amino acid residues, having two domains, namely phosphoprotein binding domain and RNA-binding domain (aa 289-352) and VSV N contains 422 amino acids. There were several structural differ-

Submission Date : 05-08-15
Revised Date : 28-08-15
Accepted Date : 13-02-16

DOI: 10.5530/ijper.50.2.29

Correspondence Address
Mr. Muthukumar Balamurugan M. Tech
Research and Development, Rasa Life Science Informatics, Pune, Maharastra-411016, INDIA.
Phone no: +91-9944776463
Email: muthubioinfotech@gmail.com

www.ijper.org

ence observed between RV and VSV nucleoprotein.[7] The one main component was identified between RV and VSV in the nucleoprotein is the accomplishing phosphorylation, where RV N gets phosphorylation but VSV N does not.[8] N protein alone non-specifically encapsidates RNA, but it needs P protein for specific genomic RNA binding and when N-P complex binds to viral genomic RNA, RV N become phosphorylated due to conformational changes.[9] When N protein binds to phosphoprotein, it introduce conformational changes in N-RNA complex which makes serine 389 residue to get phosphorylated.[6] The RV N phosphorylation makes negative repulsion between phosphoserine N and RNA causes weaker interaction and viral genome will bind to L polymerase for another cycle of transcription.[6,9] The mutational study showed when serine 389 residue gets mutated to neutral amino acids such as alanine and glycine, transcription and replication were 10 fold reduced. Increased transcription and replication were also noticed in case of acidic group amino acids such as aspartic and glutamic amino acids. This observed study further infer that the negative charge of phosphate group and structure of serine were responsible for phosphorylation of nucleoprotein that makes replication and transcription possible. Unphosphorylated N followed decrease activity of replication in minigenome system.[6] Thus phosphorylation site of serine residue is a modulating factor for transcription and replication in RV N.[6,8-9] Consequence of replication produced positive sense RNA as an intermediate, which aid as a template for synthesis negative sense viral RNA. RV N deletion–mutant study indicated that intact N preferentially encapsidates positive-strand leader RNA.[10] Nucleoprotein also known to avoid the host immune response by inhibiting Retinoic acid Inducible Gene-I-mediated antiviral production. These records point up nucleoprotein could be a potential drug target for inhibiting RV replication. In this work, Structure- based pharmacophore approach has been taken for identifying the novel ligands for Rabies nucleoprotein. The amino acid residue serine 389 was considered as an active site of the protein. Molecular docking study of Nucleoprotein with ligands derived from virtual screening divulges that ZINC01530604 and ZINC01530605 ligands molecules could be the potential lead molecule against phosphorylation site of rabies nucleoprotein.

MATERIALS AND METHODS

Sequence and Structure Extraction

The query sequence of nucleoprotein (strain SAD B19) was retrieved from Uniprot database and the structure of Nucleoprotein 2GTT was extracted from RSCB PDB in the resolution of 3.5Å.[11] GOR IV, a protein secondary structure prediction server based on information theory to formulate conformation.[12] GOR IV was used to predict secondary structure information of protein. The conserved regions of protein were detected using Pfam database.

Protein Remodeling and Loop refinement

The Accelrys Discover Studio Visualizer 3.5 was used to visualize the PDB ID 2GTT revealed that the ring like structure has 11 N protein molecules encapsidating a single strand RNA.[7] The structure of nucleoprotein is shown in Figure 1. The residues from 373-397 were found to be missed in the PDB molecule. Hence, protein remodeling was done using Phyre 2, a web portal for protein modeling, prediction and analysis server.[13] The torsion angle or dihedral angles of amino acid defines the backbone of protein. The distribution of torsion angles of modeled protein were viewed by Ramachandran plot using RAMPAGE. The loop refinement have done to minimize the loop structure in modeled protein was refined by Modeller 9.13 software package.

Energy Minimization

The molecular dynamics study was performed for analyzing stable conformation of modeled nucleoprotein. Gromacs 5.0 was used to minimize the energy minimization and simulation for molecular dynamics. Optimized Potential for Liquid Simulation (OPLS) force field and steepest descent algorithm were performed for energy minimization. The potential energy, RMSD and stability of protein were analyzed in this study.

Structure-based Pharmacophore method

Since there is no efficient drug are yet to be reported to inhibit RV entry or replication. Structure-based pharmacophore, will be an ideal method to identify the novel ligands, it is known to utilize the binding site of protein.[14] Ligand Scout 3.1 was used to generate pharmacophore features. Zinc Pharmar, free pharmacophore search database was used to generate the ligand molecule for pharmacophores from the database.[15]

Molecular Docking

Energy minimized protein and ligand molecules generated from Zinc Pharmar have been taken for molecular docking in Auto Dock Vina. Vina requires confirmation file of xyz coordinates for atomic position of amino acid where ligand will interact with macromolecule.[16]

Table 1: Secondary structure prediction analysis of SAD B 19 viral strain

Structural elements	Number of residues	% of occurrence
Helix		
Alpha helix	121	26.89
3^{10} helix	0	0
Pi helix	0	0
Strand		
Beta bridge	0	0
Extended strand	95	21.11
Coil		
Beta turn	0	0
Bend region	0	0
Random coil	234	52

Table 2: Loop docking result of Nucleoprotein with ZINC01530604 and ZINC01530605

Zinc ID	Ligand Molecules	Binding Affinity (kcal/mol)	Interaction
ZINC01530604	Molecule 1	-8.0	SER389: HG
	Molecule 2	-8.0	PHE355:HN,THR377:HN
	Molecule 3	-7.8	SER389:HN
	Molecule 4	-8.1	SER389:HG
	Molecule 5	-8.2	SER389:HG,PHE355:HG
	Molecule 6	-8.0	SER389:HG
	Molecule 7	-8.1	SER389:HG
	Molecule 8	-7.3	PHE 355:HG
	Molecule 9	-7.3	SER389:HG,PHE355:HG
	Molecule 10	-6.4	SER389:HN
ZINC01530605	Molecule 1	-7.4	SER389:HN
	Molecule 2	-7.3	SER389:HG
	Molecule 3	-8.0	PHE355:HN,THR377:HN
	Molecule 4	-8.2	LYS 396
	Molecule 5	-7.3	SER389:HN,SER389:HG
	Molecule 6	-7.2	PHE355:HN,SER389:HG
	Molecule 7	-7.3	SER89:HG
	Molecule 8	-8.2	PHE355:HN,SER389:HG
	Molecule 9	-7.3	SER389:HG
	Molecule 10	-7.3	SER389:HN

In this study we considered SER 389 amino acid as an active site of protein where ligand would interact with nucleoprotein. Vina scores nine conformations of protein-ligand interaction.

RESULT AND DISCUSSION

Secondary structure prediction

Secondary structure analysis of the sequence of viral strain SAD B 19 showed that more than 50% of the sequences were found in coil region then in helix or strand. It indicates the 3D structure of protein would have more loop region as mentioned in Table 1.

Model structure and Validation

The modeled protein showing structural elements as predicted by GOR IV was observed with missing residues which were not found in PDB is shown in Figure 2. Ramachandran plot was used for validation of protein model with plot statistics is shown in Figure 3. Loop refinement was further made for outlier region were observed in plot statistics inference that 97.4% of residues were found in favored and allowed region.

Molecular dynamics simulation

The potential energy of the modeled protein were found to be gradually decrease in energy from -1.72e +

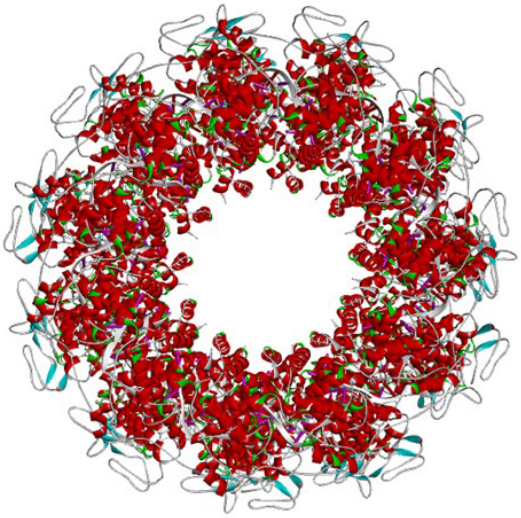


Figure 1: PDB ID 2GTT showing 11N protein chains encapsidating single strand RNA

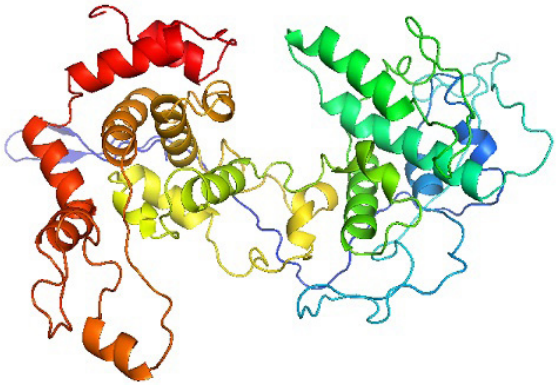
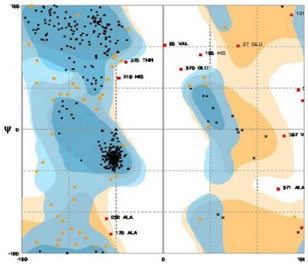


Figure 2: The 3D model of nucleoprotein with missing residues modeled by Phyre 2 server



Region	Number of Residues	Observed percentage
Favored	386	86.2
Allowed	50	11.2
Outlier	12	2.7

Figure 3: Ramachandran Plot (left) for modeled protein with plot statistics (right)

06 to -3e + 06 energy about 2000 ps inferred that model is energetically stable as plotted in Figure 4.

Pharmacophore generation

Pharmacophore were generated from protein molecule by ligand scout 3.12. It generated 1 positive (blue) and 1 negative ionizable group (red),4 hydrogen bond acceptor

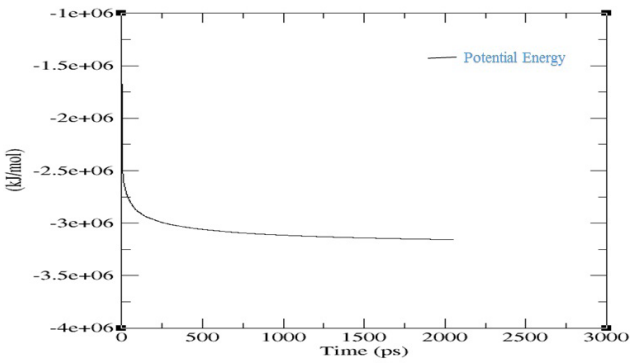


Figure 4: Potential Energy plot shows stable energy conformation for modeled nucleoprotein

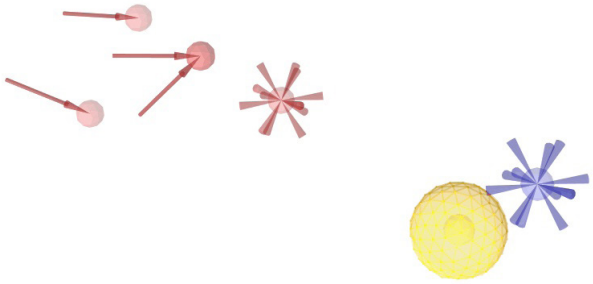


Figure 5: Pharmacophore features generated for modeled nucleoprotein

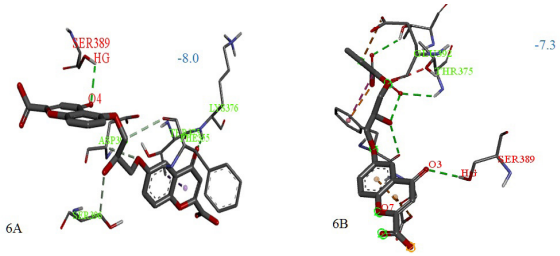


Figure 6: ZINC10530604 O4 atom showing interaction with HG atom of SER 389(red) with binding energy (blue) is shown in 6A and ZINC10530605 O3 atom showing interaction with HG atom of SER 389(red) with binding energy (blue) of nucleoprotein shown in 6B

(red arrow) and 1 hydrophobic interaction (yellow) for modeled nucleoprotein represented in Figure 5.

Binding Interaction

Initially 873 molecules were extracted having similar pharmacophore features from Zinc Pharmer. After removing the duplicates from original dataset, 246 non reductant dataset of ligand molecules were found to be unique and they were used further for docking study. The molecular docking was performed to identify the interaction between ligands and the modeled protein. On analyzing binding energy showed 12 molecules had

reasonable binding energy with Serine 389 residue. To validate the docking result, loop dock experiment was ensued. In loop docking, some of the selective ligands with good binding energy were chosen and by increasing it conformation to 20 and same ligand made to be occur for 10 times. The ligand molecules of ZINC01530604 and ZINC01530605 were found to bind with least binding energy with HN and HG of SER389 of RV N. The ligands and its binding potential is tabulated in Table 2. Molecule 1 of ZINC10530604 ligand has -8.0 binding energy with SER389 and molecule 2 of ZINC10530605 showing -7.3 binding energy with SER389 of nucleoprotein are shown in Figure 6.

CONCLUSION

Rabies, a highly neurotropic virus have ability to cause paralytic coma following to death. On literature Review, Nucleoprotein and Phosphoprotein were found to be major drug targets against rabies virus. Nucleoprotein, plays a major role in virus transcription and replication and emphasis has been given to interfere with these activities as a drug target for this present work. Nucleoprotein serine 389 residue is a major phosphorylation site, we considered the SER 389 is an active site for docking study and pharmacophore were derived from peptide of amino acids between 386-392 residues. Viral protein gets phosphorylation by means of either own viral associated protein or host cellular kinases. Caseinkinase (Ck) pathway, the host mechanism by which SER 389 gets phosphorylated in rabies nucleoprotein. The ligands such as ZINC10530604 and ZINC10530605 compounds showing favorable affinity interaction with modeled and simulated nucleoprotein and it can able to prevent the phosphorylation and thereby inhibit or reduce the transcription and replication of rabies nucleoprotein. Casein kinase inhibitor could block the transcription process but it will affect the host mechanism process. RV N phosphorylation prevention is more safety measure than host Ck block. Anyhowfuture *in vivo* experimental studies are required to provide a new insight to validate the results for such ligand molecules.

ACKNOWLEDGEMENT

The author have none to declare.

CONFLICT OF INTEREST

The author have no conflict of interest.

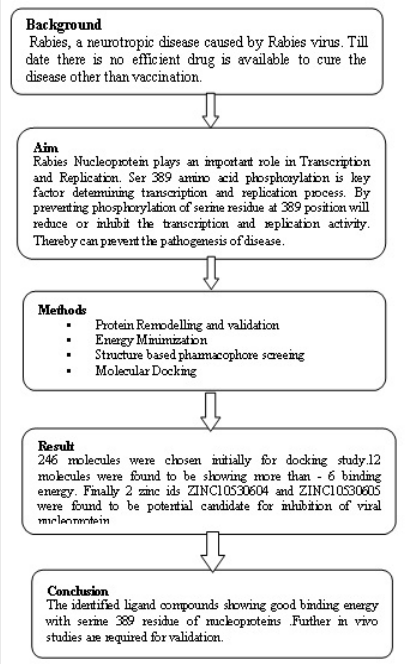
REFERENCES

1. Albertini AAV, Blondel RWH, Bolondel D. Rabies Virus Transcription and Replication. Advances in Virus Research. 2011;79(1):2-16.
2. Knobel DL, Cleaveland S, Coleman PG. Re-evaluating the burden of rabies Africa and Asia. Bull World Health Organ. 2005; 83(5):360-8.
3. Oksayan S, Ito N, Moseley G, Blondel D. Subcellular Trafficking in Rhabdovirus Infection and Immune Evasion: A Novel Target for Therapeutics. Infectious Disorders-Drug Targets. 2012;12(1):38-58.
4. Wiltzer L, Okada K, Yamaoka S, Larrous F, Kuusisto HV, Sugiyama M. Interaction of Rabies Virus P-Protein with STAT Proteins is Critical to Lethal Rabies Disease. Journal of Infectious Diseases. 2014;209(11):1744-53. [PMID:24367042]
5. Graham SC, Assenberg R, Delmas O, Verma A, Gholami A. Rhabdovirus Matrix Protein Structures Reveal a Novel Mode of Self Association. PLoS Pathogens. 2008;4(12):e1000251. doi:10.1371/journal.ppat.1000251.
6. Wu X, Gong X, Foley HD, Schnell MJ, Fu ZF. Both Viral Transcription and Replication Are Reduced when the Rabies Virus Nucleoprotein Is Not Phosphorylated. Journal of Virology. 2002;76(9):4153-61. DOI: 10.1128/JVI.76.9.4153-4161.2002.
7. Luo M, Green TJ, Zhang X, Tsao J, Qiu S. Conserved characteristics of the rhabdovirus nucleoprotein.Virus Res. 2007;129(2):246-51.
8. Yang J, Koprowski H, Dietzschold B, Fu ZF. Phosphorylation of Rabies Virus Nucleoprotein Regulates Viral RNA Transcription and Replication by Modulating Leader RNA Encapsidation. Journal of Virology. 1999;73(2):1661-4.
9. Liu P, Yang J, Wuand X, Fu ZF. Interactions amongst rabies virus nucleoprotein, phosphoprotein and genomic RNA in virus-infected and transfected cells. Journal of General Virology. 2004;85(12):3725-34.
10. Yang J, Hooper DC, Wunner W.H, Koprowski H, Dietzschold B, Fu Z.F. The Specificity of Rabies Virus RNA Encapsidation by Nucleoprotein. VIROLOGY. 1998;242(1):107-17.
11. Albertini AAV, Wernimont AK, Muziol T, Ravelli RB, Clapier CR, Schoehn G, *et al*. Crystal structure of the rabies virus nucleoprotein-RNA complex. Science. 2006;313(5738):360-3. [PubMed: 16778023].
12. Garnier J, Gibrat JF, Robson B. GOR secondary structure prediction method version IV. Methods in Enzymology. 1996; R.F. Doolittle Ed.266:540-53.
13. Kelley LA. The Phyre2 web portal for protein modeling, prediction and analysis. Nature Protocols. 2015;10(6):845-58.
14. Gaurav A, Gautam V. Structure-based three-dimensional pharmacophores as an alternative to traditional methodologies. Journal of Receptor Ligand and Channel Research. 2014;7:27-38.
15. Koes DR, Camacho CJ. Zincpharmer: pharmacophore search of the ZINC database. Nucleic Acids Research. 2012;40(w1): 409-14.
16. Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function efficient optimization and multithreading. J Comput Chem. 2010;31(2):455-61.

SUMMARY

- Prior vaccination to Rabies virus and/or immediate vaccination before onset of clinical symptoms are available treatment for rabies fever.
- In this Present study, we identified 2 Zinc ID ZINC10530604 and ZINC10530605 from Zinc pharmar database showing highest binding energy with ser389 residue of nucleoprotein.

PICTORIAL ABSTRACT



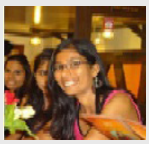
About Authors

Muthukumar Balamurugan: Is a Research Trainee in RASA Life Science Informatics. His research area focused on CADD, Molecular Dynamics.

Aril Bhoomigowda: Is a Research Trainee in RASA Life Science Informatics. Her research area focused on CADD, cheminformatics.

Tanveepardeshi: Is a Junior Application Scientist in RASA Life Science Informatics. She specializes in research and training in the area of Bioinformatics, Chemo informatics, CADD, Molecular Dynamics, NGS Data Analysis.

Sameer Choudhary: Is the CSO and Co-founder at RASA Life Science Informatics. His expertise includes machine learning and data mining, scientific workflow development, product development, offshore services and Management.



Shama Mujawer: Is working as a Junior Application Scientist in RASA Life Science Informatics. She specializes in research and training in the area of Bioinformatics, Chemo informatics, CADD, Molecular Dynamics, NGS Data Analysis.