Proceedings of the 3rd International Conference on Biological Engineering and Medical Science
DOI: 10.54254/2753-8818/21/20230812

Rabies' proteins' functions and future directions

Xintian Lyu

Suzhou North America School, Suzhou, China

1811000714@mail.sit.edu.cn

Abstract. The rabies virus, which mostly originates from the bite of a sick dog or cat and spreads, kills tens of thousands of people each year, and it is super lethal, with almost no one surviving the infection, s so, everyone is afraid of it. The main body of the rabies virus is bullet-shaped, and each combination of proteins in it is an innate destroyer. From the destruction of cells by the G and M proteins to the transcription of the P and L proteins in concert with each other, the precision, and division of labor of the rabies virus are obvious. Of these, the L protein is the key to the operation of everything, and this paper will explore the destruction of each protein, thereby stopping the spread of RABV. In particular, the possibility that the L protein, unlike rabies vaccines currently on the market, might be a future drug design idea by destroying the structure of RABV after the virus enters the cell. Keywords: Rabies, RABV, Protein Matching, L protein.

Keywords: rabies, RABV, protein matching, L protein.

1. Introduction

Rabies is a harmful neurological disease caused by rabies virus (RABV) infection. Rabies virus causes acute zoonotic infections, mostly in carnivores such as dogs, wolves, and cats. Humans are mostly infected by diseased animals' bites. Clinical manifestations include a characteristic fear of water, wind, pharyngeal muscle spasms, and progressive paralysis. There is almost no chance of survival if clinical signs appear [1]. Rabies vaccination is given after being bitten by an animal that carries the rabies virus. This event is done so that the antibodies are present in the body for a long time. This way, when the actual disease-causing rabies virus appears, the antibodies can quickly destroy it before it can multiply. There is no effective treatment for rabies, but exploring the amino acid sequence of the L protein might be helpful.

The rabies virus is transmitted to humans and animals through the saliva of a sick animal and then enters the infected body through a wound. Invasion of the neuromuscular junction followed by interaction with nicotinic acetylcholine receptor proteins spreads the infection to other peripheral nerves. Subsequently, the extended infected nerve spreads to the central nervous system. This event results in damage to neurons in the brainstem and cerebellum. In the gray matter of the brain, the viral nucleoshell is expelled from the endosome and transported along microtubules to form new virosomes or secondary viral factories [2]. It contains five proteins: polymerase (L), matrix protein (M), glycoprotein (G), and nucleoprotein (N). In figure 1, the genes for these five proteins line up after the 58 nt leader sequence in front of Q, in the order 3′-N-P-M-G-L-5′. After these are followed by the 57-70 nt trailer [3]. The RABV divided into two parts: the surfing envelope and the RNP. Nucleoprotein. The surrounding envelope is composed of knob-like spikes of glycoprotein, while the RNP contains Nucleoprotein [4].

DOI: 10.54254/2753-8818/21/20230812

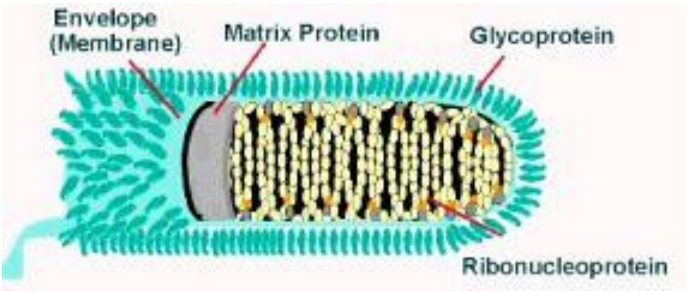

Figure 1. shows the rabies virus, which is coated in spike-like glycoprotein peplomers that are 10 nm in size [4].

Glycoprotein (G) catalyzes and causes membrane fusion, which enables the virus to enter the groundplasm and cause infection. Succeeding full entrance into the cell, the P protein helps the L polymerase transcribe the genome of the virus to create circulating viral proteins [5]. The L protein produces five strands of mRNA and one strand of positive RNA from the first negative-stripped RNA. These five mRNA strands are translated into the relevant proteins on ribosomes that are floating in the cytoplasm. Post-translational modifications are necessary for some proteins. When there are sufficient viral proteins present, viral polymerase makes negative-stranded RNA from positive-stranded RNA templates. These averse strands will go to the inner cell membrane in a combination with the protein. The newly created the pathogenic particle's outer envelope is made up of a G protein that is incorporated in the membrane and surrounds the protein complex. When the viruses reach a certain number of replications, they start to bind to acetylcholine receptors at the neuromuscular junction and cross the axon of the nerve cell by retrograde transport. In this section, their P protein interacts with proteins in the cytoplasm of the nerve cell, causing the virus to swiftly enter the brain's nervous system once it reaches the cell body, replicating in motor neurons, and eventually reaching the brain. L protein is essential to the way that RABV is transmitted. It is the biggest structural protein of the RABV and necessary for the synthesis of every viral protein, including the virus itself.

The negative-stripped RNA is converted into a positive strand of RNA by the L protein from the original negative-stripped RNA. On free ribosomes in the cytoplasm, these five mRNA strands are translated into the necessary proteins. Some proteins require post-translational changes. Viral polymerase converts positive-stranded RNA templates into negative-stranded RNA when there are enough viral proteins present.

2. About G protein

Treatment for RABV generally begins with G protein. In figure 2, there is a type I glycosylated protein and a trimer consisting of 524 amino acids. This polypeptide contains 524 amino acids, including 19 amino acids in the signal sequence. Arginine at locus 333 is crucial in the virulence of RABV. This is associated with neuroinvasiveness and the ability to spread across synapses, enabling the virus to spread faster in the nervous system. The glycoprotein is a significant factor that determines whether or not RABV can enter and infect [6]. In order to induce endocytosis of the virion, the virus binds to the host cell's receptor via the G protein. The acidic pH in the endosome causes conformational changes in the glycoprotein trimer, which results in membrane fusion. Low pH induces different conformational states in elastovirus glycoprotein during membrane fusion. Blister virus glycoproteins can be classified by their structural features as pre-fusion, early intermediate, late intermediate, and post-fusion. While RABV virus glycoproteins usually exist as trimeric spikes on their surface, the RABV virus glycoproteins are mainly trimeric spikes. In solution, however, it mainly exists as a monomer, not as a trimer [7]. Various in vitro experiments have shown that the muscular form of the nicotinic acetylcholine receptor (nAChR) binds glycoprotein, facilitating rabies virus entry [8]. That is the reason why the investigation wants to study glycoprotein because if it can cut the beginning of the infection.

DOI: 10.54254/2753-8818/21/20230812

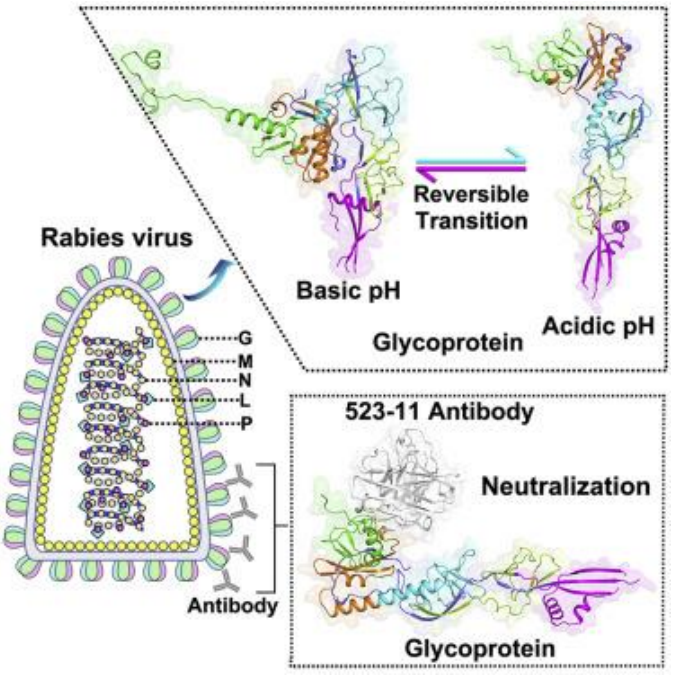

Figure 2. Analysis of the rabies virus glycoprotein's structure [7].

3. About M protein

The M protein is necessary for viral assembly, acts as a matrix and is a major component of the entire virion. As a pathogenicity determinant, it determines the virus's assembly and germination. Cell membrane disruption occurs due to the amino acid at position 95 of the M protein [8]. The previously mentioned viruses generate nucleocapsids when they enter the gray matter of the brain. this ability to bind and possibly coalesce nucleocapsids exists in the M protein. It can also help the G protein to enter [9]. While M proteins interact with RNP and G proteins and play an important role in recruiting RNP to host cell membranes and in outgrowing viral particles, they also interact with other proteins. It is also possible for M proteins to act synergistically with the JAK-STAT pathway in the absence of G proteins to regulate the pathway [10].

4. About N protein

Nucleocapsid proteins are proteins. The N protein always encases genomic RNA. N proteins congregate with ribonucleic acid to inhibit nucleases [11]. Both N-RNA complexes include single-stranded RNA with random sequences and are organized as loop assemblies. It aids in wrapping and folding virus particles. It has been established that the N protein's pathogenicity determinants, amino acids 273 and 294, block the generation of IFN, which in turn affects the ability of viruses to enter the brain. This shows that by enabling successful viral transmission, the RABV N protein contributes significantly to avoiding the brain's instinctive immune system reaction [11]. N proteins are equally as important as the G protein in enabling membrane fusion to avoid innate immune system components during vivo. N proteins play a crucial role. They influence the ability of in vivo proteins to inhibit the host antiviral response. Mainly, by facilitating efficient virus transmission [9].

5. About P&L protein

It was discovered that the NPYNE sequence is necessary for the interactions between the L and P proteins, and this region is indicated in light orange and pink in the figure 3. The binding domain of this L-protein sequence is at the C-terminus of P-protein-L. Because L proteins require the interaction with their important cofactor P proteins, the activity of L proteins as RdRp all depends on this region [17].

DOI: 10.54254/2753-8818/21/20230812

Scientists find that the L protein is the largest structural protein, consisting of 2127-2142 amino acids, and is an enzyme complex of approximately 244 kDa that is fundamental for producing all RABV proteins, including its own, because it functions as a dependent RNA polymerase (RdRp). In the figure 4, L protein can be classified into three macromolecular structures: Glycoprotein, RVA122 Fab Light Chain, and RVA122 Fab Heavy Chain. The L protein has a small receptor molecule called NAG [12]. Since the L protein is associated with all other RABV genes, inhibiting the L protein from it could be considered. L proteins are of significance for viral binding, replication as well as transcription [10]. In the cytoplasm, negative-stranded RNA is translated into five strands of mRNA and one strand of positive RNA. By free ribosomes located in the cytoplasm, all five of these mRNA strands are translated into proteins. It is noteworthy that in this process, the interplay of L and P proteins is required for completion. Despite the lack of information on the three-dimensional structure of the L protein, RABV's L protein is significantly similar to the fully characterized L protein of VSV in amino acid sequence, structural domain structure, and enzymatic function. This was then compared with the structure determined by cryo-electron microscopy of the VSV L protein. Although the L protein binding site could not be determined, it was seen that if started from the RABV P protein, it stimulated the initiation of transcription, and the extension mediated by the L protein [13].

The P protein is a regulatory and non-catalytic polymerase cofactor that aids in viral transcription and replication [14]. The endosomal structure requires the development of dimerized structural domains of P proteins. This stretch, which is mediated by the P and L proteins, is found in the residues of the P protein. A significant loss in the carboxy-terminal region of P does not affect its ability to interact with the L protein. The trials and the outcomes do not necessarily prove that L proteins need P proteins. However, the P protein and L protein cannot be separated. The former also binds to the soluble protein N to prevent it from wrapping around non-genomic RNA and remains on the N-RNA template to keep the RNA polymerase L stable. The P protein contains two N-binding sites. As a result, the P protein's N and L protein binding sites do not overlap. P proteins can behave as replication factors or transcription factors depending on whether they are paired with L proteins or N proteins. The P protein binds to the N protein, which in turn directs the G protein, so long as it can block the L protein while also destroying the P protein, which in turn destroys the N and G proteins [15].

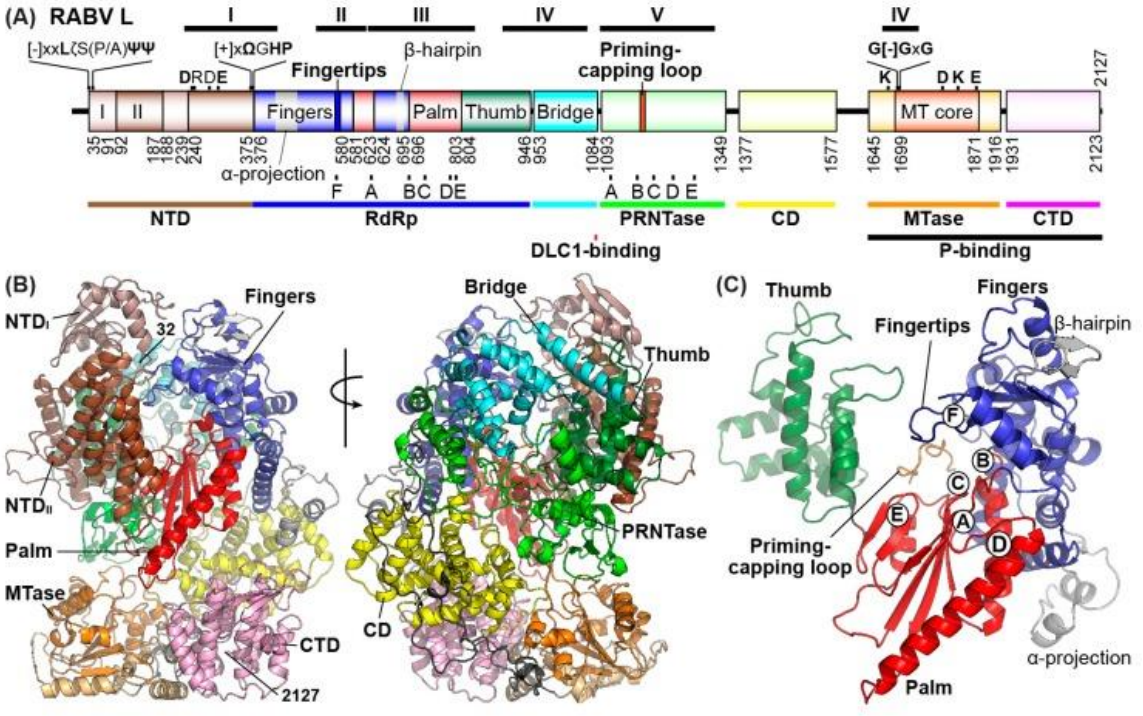

Figure 3. The L protein of RABV [13].

DOI: 10.54254/2753-8818/21/20230812

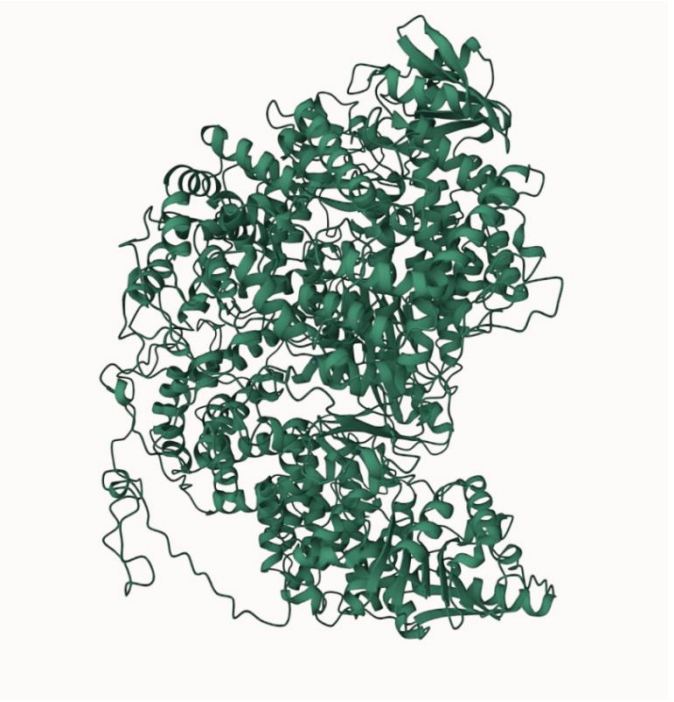

Figure 4. Structure guessing of L protein in RABV [16].

The graphic above displays the expected L protein sites for the RABV based on VSV. The numbers indicate the start and end points of the structural and sub-structural domains of amino acid residues. The picture shows the preserved region (I-VI), the RdRp pattern (A-F), and the PRNTase pattern (A-E). The L gene-deficient RABV (Nishi-L/Nluc) was developed by Kento Nakagawa et al. through reverse genetics. NanoLuc luciferase was used to transfect cultured neuroblastoma cells in order to create L protein. To determine the functional importance of the highly conserved L protein region between positions 1914 and 1933, back-complementation was employed with mutant L proteins. It was discovered that the NPYNE sequence at positions 1929 to 1933 is necessary for the interactions between the L and P proteins, which is also shown on the graph in light orange and pink. The binding domain is at the C-terminus of P protein-L. The interaction of the L protein with the P protein, and therefore the L protein's activity as RdRp, depend on this area [17].

6. Conclusion

The proteins of rabies virus have a very close cooperation: the G protein, assisted by the N protein, serves as the main destructive protein structure of rabies. The main body of the virus is the M protein, and the critical viral transformation depends on the cooperation of L and P proteins. The important sites of cooperation between the P and L proteins are at amino acids 1929 to 1933 of the L protein of RABV, which can be designed as a drug target for direct destruction not in the external tissues but in the interior. In other words, starting with the L protein may contribute to new advancement in treating rabies in the future.

References

[1] "Rabies." https://www.who.int/news-room/fact-sheets/detail/rabies
[2] Davis, Benjamin M, et al. "Everything You Always Wanted to Know about Rabies Virus (but Were Afraid to Ask)." Ann. Rev. Virol.. 2015,
[3] "Genus: Lyssavirus."https://ictv.global/report/chapter/rhabdoviridae/rhabdoviridae/lyssavirus.
[4] "What Is Rabies?" https://www.cdc.gov/rabies/about.html?CDC_AA_refVal=https%3A%2F%2Fwww.cdc.gov%2Frabies%2Ftransmission%2Fvirus.html.
[5] Albertini, A. A. V., et al. "Structural Aspects of Rabies Virus Replication" Cell. Mol. Life Sci. 65 282, (2007).

DOI: 10.54254/2753-8818/21/20230812

[6] Ben Khalifa, Youcef, et al. “The Matrix Protein of Rabies Virus Binds to RELAP43 to Modulate NF-KB-Dependent Gene Expression Related to Innate Immunity.” Sci. Rep., 21 (2016).
[7] Yang, Fanli. “Structural Analysis of Rabies Virus Glycoprotein Reveals PH-Dependent Conformational Changes and Interactions with a Neutralizing Antibody”, Cell host & microbe, 27(3): 441-453 (2020).
[8] “Uniprot Website Fallback Message.” https://www.uniprot.org/uniprotkb/P03524/entry.
[9] Mebatsion, T, et al. “Matrix Protein of Rabies Virus Is Responsible for the Assembly and Budding of Bullet-Shaped Particles and Interacts with the Transmembrane Spike Glycoprotein G.” J. Virol. 73, 242, (1999)
[10] Yin, Juanbin, et al. “Research Advances on the Interactions between Rabies Virus Structural Proteins and Host Target Cells: Accrued Knowledge from the Application of Reverse Genetics Systems.” Viruses, 16, 2021,
[11] “Uniprot Website Fallback Message.” https://www.uniprot.org/uniprotkb/P16285/entry.
[12] “Uniprot Website Fallback Message.” https://www.uniprot.org/uniprotkb/A3F5L9/entry.
[13] Ogino, Tomoaki, and Todd J Green. “Transcriptional Control and Mrna Capping by the GDP Polyribonucleotidyltransferase Domain of the Rabies Virus Large Protein.” Viruses, 1 (2019),
[14] “Uniprot Website Fallback Message.” UniProt, https://www.uniprot.org/uniprotkb/P22363/entry.
[15] Chenik, M, et al. “Mapping the Interacting Domains between the Rabies Virus Polymerase and Phosphoprotein.” J. Virol. (1998)
[16] Green, Todd “Model of the Rabies Virus L Protein.” ModelArchive, (2020)
[17] Nakagawa, Kento, et al. “Molecular Function Analysis of Rabies Virus RNA Polymerase L Protein by Using an L Gene-Deficient Virus.” J. Virol., U.S. National Library of Medicine, 2017.