Structural and Physico-chemical Analysis of the Matrix Protein of Nipah Virus (NiV): Implications for Drug Design

Sohpal Vipan K.[1]* and Gill Gurleen[2]
1. S.B.S State University, Gurdaspur, Punjab, INDIA
2. I.K Gujral Punjab Technical University, Jalandhar, Punjab, INDIA
*vksohpal@sbssugsp.ac.in

Abstract

The Nipah virus, a zoonotic pathogen with high mortality rates, poses a significant public health concern. This research focuses comprehensive analysis of its structural and physico-chemical properties, this study aims to contribute valuable insights for drug design and therapeutic interventions. The matrix protein's amino acid sequence was accessed and subjected to physicochemical characterization, revealing stable features with an instability index of 29.53 and a high aliphatic index of 90.26. Secondary structure prediction using SOPMA indicated a prevalence of random coils (51.70%), while PHYRE 2 provided additional insights into disorder prediction.

Homology modeling through Swiss Model generated a three-dimensional structure, exhibiting good stereochemical quality validated by a Ramachandran plot. Approximately 92.2% of residues were located in favourable regions, indicating the reliability of the predicted model.

Keywords: Nipah Virus (NiV), Physicochemical, Secondary structure, Homology modeling, Ramachandran plot.

Introduction

The Nipah virus, a member of the Paramyxoviridae family (Henipavirus genus), exhibits a unique structure crucial for infection and pathogenesis. Its genome comprises of RNA, enclosed in an envelope with viral glycoproteins (G and F) facilitating cell entry. The matrix protein provides structural support and the ribonucleoprotein complex aids in replication. The virus emerged from fruit bats, causing fatal diseases in humans and animals. Henipaviruses, including Nipah, are classified as BSL4 pathogens due to their zoonotic potential[3].

The Nipah virus, a negative-stranded, ssRNA virus, features six principal genes[7]. Recent outbreaks in animals and humans underscore its threat, with symptoms ranging from fever to fatal central nervous system infections[8]. NiV-M, a main structural protein and its membrane association, protein-protein interactions play vital roles in viral envelope formation and budding. NiV-M's disruption of host cell functions contributes to the virus's pathogenicity. Understanding NiV-M's molecular mechanisms, achieved through bioinformatics tools, offers insights into its structure, physicochemical properties and impact on host cells. This study aims to perform *in silico* analysis and 3D structure prediction of NiV-M, an unreported structure. Utilizing tools like PHYRE2 and Swiss Model will contribute to protein-based drug and vaccine design, potentially minimizing Nipah virus infections[7,10].

Review of Literature

NiV was identified as a paramyxovirus originating from bats of the genus Pteropus, causing deadly diseases in individuals[4]. NiV belonged to the negative-stranded, ssRNA genome category, constituting a new genus, Henipavirus, within the Paramyxoviridae family[3]. The role of bioinformatics in addressing biotechnological challenges was attributed to the availability of sequential and structural information, making *in silico* approaches more time and cost-effective than experimental methods[7]. Although the tertiary structure of NiV matrix protein (NiV-M) was unknown, crystal structures of M proteins from other paramyxoviruses, such as Newcastle disease virus (NDV, Avulavirus) and the distantly related respiratory syncytial virus (RSV, Pneumovirus), provided insights[10].

Despite low primary sequence homology (~20%), NDV and RSV matrix proteins shared a common architecture. Homology modeling suggested similar structures in other M proteins, emphasizing surface-exposed features driving membrane association. NiV-M formed stable dimers, associating into pseudotetrameric arrays crucial for viral particle budding. Ectopic expression of M alone facilitated budding, but during infection, M recruited and incorporated other viral structural proteins[10]. Characterization of NiV lacking the M gene underscored its pivotal role in viral stability and morphology, highlighting regulatory aspects in paramyxovirus biology. The initial NiV infection cluster emerged in September 1998 in Perak, Malaysia, followed by clusters in Negri Sembilan, primarily involving adult men in contact with swine[4].

Subsequent eruptions in Singapore and Malaysia led to interventions, including pig culling, resulting in the end of the outbreaks nearly two years later[5]. Sporadic NiV outbreaks in South and Southeast Asia exhibited variations in clinical presentation, case fatality rate and transmission modes. A new NiV strain in Bangladesh and India (2000–2001) showcased acute severe respiratory symptoms, raising concerns about human-to-human transmission[8]. Consumption of raw date juice emerged as a common

behavior among infected individuals, emphasizing NiV transmission via contaminated date palm sap[6].

NiV outbreaks in the Philippines (2014) associated with horse slaughter and consumption of horse meat were reported. Although NiV outbreaks had not recurred in Malaysia or Singapore since 1998–1999, annual outbreaks persisted in Bangladesh, with sporadic occurrences in India[1].

Clinical and autopsy findings during the initial NiV outbreak in Malaysia and Singapore (1998–1999) revealed rapid progression, with symptoms largely localized to the central nervous system (CNS)[4]. Pathological features included vasculitis and necrotic plaques, emphasizing the virus's impact on multiple organs[2]. The strain responsible for recurrent outbreaks in Bangladesh (NiV-B) differed from the Malaysian/Singaporean strain (NiV-M), showing distinct clinical and epidemiological characteristics. NiV-B was linked to respiratory symptoms, human-to-human transmission and higher case fatality rates, suggesting viral virulence factors contribute to mortality[9].

Material and Methods

The amino acid sequence (Q9IK90) of NiV matrix protein, comprising of 352 amino acids, was retrieved from the NCBI sequence database. Expasy's ProtParam server evaluates the physico-chemical attributes of NiV matrix protein that significantly impact the virus life cycle, including replication and pathogenesis properties. Using PHYRE 2 and SOPMA tools, secondary structure prediction of matrix proteins in Nipah virus is conducted, focusing on folding patterns and functional insights. Homology modeling via Swiss Model generates the 3-D structure of the matrix protein. The structural models undergo quality assessment through Ramchandran plot analysis and evaluation using the PDBsum tool, ensuring reliability for further investigations.

Results and Discussion

The amino acid sequence (Q9IK90.1) of the Matrix protein from NiV Henipavirus was sourced from the NCBI database. Utilizing the 352-amino-acid-long sequence, the tertiary structure of protein Q9IK90.1 was modeled. Supplementary details about the protein (Q9IK90.1) are available in table 1.

The FASTA-formatted amino acid sequence Q9IK90.1, extracted from Nipah Henipavirus, served as the query sequence for assessing physicochemical parameters. Table 2 represents the computed physico-chemical parameters using Expasy's ProtParam tool. A stability prediction below 40 for the instability index designates a stable protein. Our analysis with an index of 29.53 indicates stability. The notably high extinction coefficient (43235) suggests a concentration of Cys, Trp and Tyr, known for UV light absorption. Aromatic amino acids (tyrosine, tryptophan, cysteine) contribute to UV absorbance at 280nm wavelength in proteins and peptides.

The aliphatic index, a positive factor for thermal stability, reflects the relative volume occupied by aliphatic side chains (alanine, valine, leucine, isoleucine). The query protein's elevated aliphatic index (90.26) implies stability across a wide temperature range. The Grand Average Hydropathy (GRAVY) value, the average hydropathy of a protein, stands at a low (-0.211) level, suggesting favourable water interaction. With 36 negatively charged residues and 48 positively charged residues, the matrix protein exhibits specific charge characteristics.

The intricate connection between protein structure and function is well-established. Secondary structural elements like helix, coil, sheet and turn play a crucial role in shaping protein function, structure and interactions. Two software tools, SOPMA (Self Optimized Prediction Method with Alignment) and Phyre 2, were employed to predict the secondary structure of the Nipah virus matrix protein.

Table 1
Protein Retrieval

Protein individualities	**Protein information**
LOCUS	MATRX_NIPAV
Amino acid	352 aa
Definition	Matrix protein [Nipah Henipavirus]
Accession	Q9IK90
Version	Q9IK90.1
Source	Henipavirus nipahense
Organism	Henipavirus nipahense
FASTA sequence	>sp|Q9IK90.1 MATRIX protein [Nipah Henipavirus] MEPDIKSISSESMEGVSDFSPSSWEHGGYLDKVEPEIDENGSMIPKYKIYTP GANERKYNNYMYLICYGFVEDVERTPETGKRKKIRTIAAYPLGVGKSASHP QDLLEELCSLKVTVRRTAGSTEKIVFGSSGPLNHLVPWKKVLTSGSIFNAVK VCRNVDQIQLDKHQALRIFFLSITKLNDSGIYMIPRTMLEFRRNNAIAFNLL VYLKIDADLSKMGIQGSLDKDGFKVASFMLHLGNFVRRAGKYYSVDYCR RKIDRMKLQFSLGSIGGLSLHIKINGVISKRLFAQMGFQKNLCFSLMDINPW LNRLTWNNSCEISRVAAVLQPSIPREFMIYDDVFIDNTGRILKG

Figure 1 displays the SOPMA results, revealing a higher prevalence of random coils (51.70%) compared to other secondary structure elements such as alpha helix (17.5%), extended strand (25.57%) and beta turns (5.68%). SOPMA utilized default parameters (window width: 17, similarity threshold: 8 and number of states: 4) for secondary structure prediction, secondary structure and disorder prediction, Phyre 2 was employed.

To forecast the 3-D structure of the matrix protein, homology modeling is employed, utilizing software such as Swiss Model. The resultant structure from Swiss Model is then visualized using Rasmol as shown in figure 2. To assess the stereochemical quality of the predicted structure generated by Swiss Model, Ramchandran map calculations are conducted using PROCHECK software as represented in figure 3.

Table 2
Computed Physicochemical Parameters

Property	Value
Number of amino acids	352
Molecular weight	39928.28
Theoretical pI	9.31
Total number of negatively charged residues (Asp + Glu)	36
Total number of positively charged residues (Arg + Lys)	48
Extinction coefficient	43235
Instability index	29.53
Aliphatic index	90.26
Grand average of hydropath city (GRAVY)	-0.211

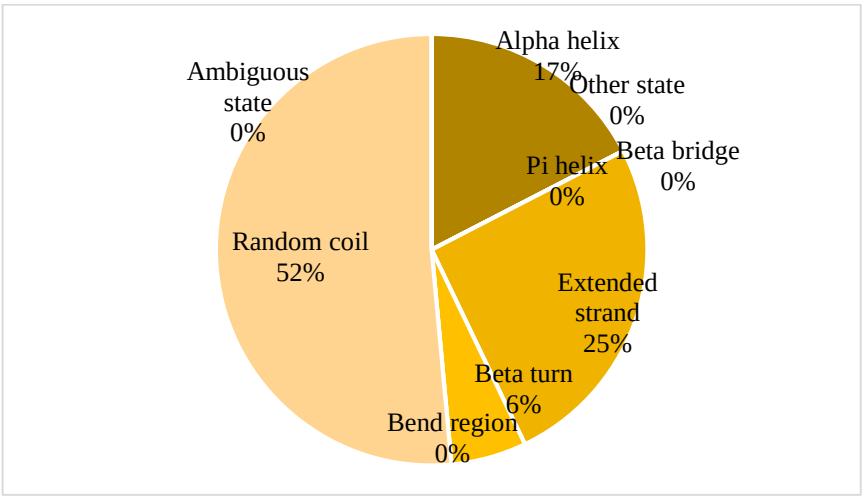

Figure 1: Secondary Structure Parameters Composition –SOPMA

Figure 2: Modeled Structure of Matrix Protein of NiV

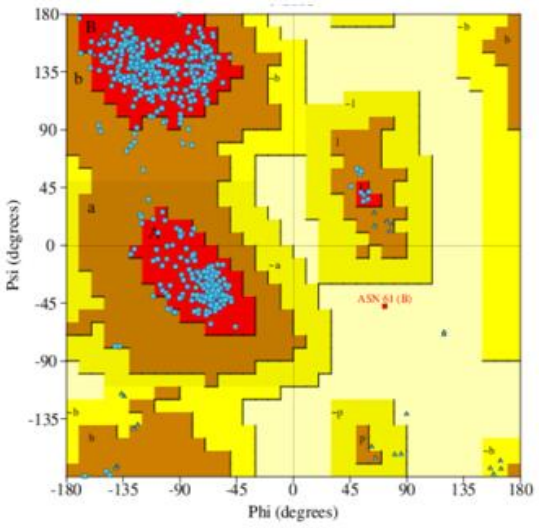

Figure 3: Ramachandran plot of Matrix Protein of NiV

Table 3
Ramachandran Plot Analysis

Parameters	No. of residues	Value %
Most favourable regions (A, B, L)	449	92.2
Additional allowed region (a, b, l, p)	37	7.6
Disallowed regions	1	0.2
Non-glycine and non-proline residues	487	100
End residues (Excl. Gly and Pro)	19	-
Glycine residues	33	-
Proline residues	20	-

A significant portion, 92.2% of residues, is located in the core regions including right-handed alpha helices (A), beta sheets (B) and left-handed alpha helices (L), as depicted by red areas in the plot, indicating the most favourable regions. 7.6% of residues are situated in additional allowed regions (a, b, l, p), illustrated in brown color. Only a minimal 0.2% of residues are found in disallowed regions depicted in table 3. These results collectively suggest a high quality of the predicted model.

Physico-chemical parameters, outlined in table 2, were computed using Expasy's ProtParam tool. Notably, the instability index of 29.53 indicated the protein's stability. The high extinction coefficient (43235) suggested a concentration of UV-absorbing amino acids, contributing to the protein's interaction with UV light. Additionally, the aliphatic index (90.26) hinted at the protein's stability across a broad temperature range. The grand average hydropathy value of 0.211 suggested favourable water interaction, while the specific charge characteristics were elucidated through the count of negatively and positively charged residues.

The secondary structure prediction, conducted using SOPMA and Phyre 2, revealed insights into the composition of helix, coil, sheet and turn elements. SOPMA results (Figure 1) indicated a prevalence of random coils (51.70%) compared to other secondary structure elements. The tertiary structure prediction through homology modeling using Swiss Model provided a modeled structure (Figure 2) of the Matrix protein, showcasing potential insights into its three-dimensional arrangement.

Conclusion

The simulative analysis of the NiV matrix protein (Q9IK90.1) unveiled essential characteristics. The protein demonstrated stability, UV-absorbing properties and a favourable interaction with water. Secondary structure predictions highlighted the prevalence of random coils, shedding light on the protein's structural composition. The 3-D model attained through homology modeling indicated a high-quality structure, with a significant portion of residues situated in favourable regions. This multifaceted analysis contributes to a deeper understanding of the physico-chemical, structural and functional aspects of the NiV matrix protein, laying the groundwork for further investigations and potential applications in drug design and development.

References

1. Banerjee S., Gupta N., Kodan P., Mittal A., Ray Y., Nischal N. and Wig N., Nipah virus disease: A rare and intractable disease, *Intractable & Rare Diseases Research*, **8(1)**, 1-8 **(2019)**

2. Chadha M.S. et al, Nipah virus-associated encephalitis outbreak, Siliguri, India, *Emerging Infectious Diseases*, **12(2)**, 235 **(2006)**

3. Gazal S. et al, Nipah and Hendra viruses: deadly zoonotic paramyxoviruses with the potential to cause the next pandemic, *Pathogens*, **11(12)**, 1419 **(2022)**

4. Hauser N., Gushiken A.C., Narayanan S., Kottilil S. and Chua J.V., Evolution of Nipah virus infection: past, present and future considerations, *Tropical Medicine and Infectious Disease*, **6(1)**, 24 **(2021)**

5. Kulkarni D.D., Tosh C., Venkatesh G. and Senthil Kumar D., Nipah virus infection: current scenario, *Indian Journal of Virology*, **24(3)**, 398-408 **(2013)**

6. Nahar N. et al, Raw sap consumption habits and its association with knowledge of Nipah virus in two endemic districts in Bangladesh, *PLoS One*, **10(11)**, e0142292 **(2015)**

7. Saikat A.S.M., Paul A.K., Dey D., Das R.C. and Das M.C., In-Silico Approaches for Molecular Characterization and Structure-Based Functional Annotation of the Matrix Protein from Nipah henipavirus, Chemistry Proceedings, **12(1)**, 21 **(2022)**

8. Shariff M., Nipah virus infection: A review, *Epidemiology & Infection*, **147**, e95 **(2019)**

9. Singh R.K. et al, Nipah virus: epidemiology, pathology, immunobiology and advances in diagnosis, vaccine designing and control strategies–a comprehensive review, *Veterinary Quarterly*, **39(1)**, 26-55 **(2019)**

10. Watkinson R.E. and Lee B., Nipah virus matrix protein: expert hacker of cellular machines, *FEBS Letters*, **590(15)**, 2494-2511 **(2016)**.

(Received 19th January 2024, Revised 17th June 2025, accepted 05th July 2025)