Iran J Public Health, Vol. 53, No.12, Dec 2024, pp.2789-2801

Original Article

Identification of CD8+ Immunogenic Peptides for Vaccine Design against Nipah Virus in Humans

Pankaj Suden, *Inderpal Singh, Chitrakshi Chopra, Malvika Mehta, Ratna Chandra, *Indu Bhushan

School of Biotechnology, Shri Mata Vaishno Devi University, Katra, India

***Corresponding Authors:** Emails: indu.bhushan@smvdu.ac.in, ipsinghbijral@gmail.com

(Received 07 Feb 2024; accepted 18 Apr 2024)

Abstract

Background: Nipah virus is a pathogenic virus of ruinous zoonotic potential with inflated rate of mortality in humans.

Methods: Considering the emerging threat of this pandemic virus, the present investigation amid to design vaccine by using the bioinformatics tools such as host and virus codon usage analysis, CD8+ peptide prediction, immunogenicity/allergenicity/toxicity, MHC-I allele binding prediction and subsequent population coverage and MHC-I-peptide docking analysis.

Results: In this study (conducted in 2022 at School of Biotechnology, Katra, India), a set of 11 peptides of the structural proteins of Nipah Virus were predicted and recognized by the set of MHC-I alleles that are expressed in 92% of the global human population.

Conclusion: The strong interactions between these peptides and the MHC-I protein suggest them as strong peptide candidates for the development of vaccine against Nipah Virus.

Keywords: Zoonotic virus; Nipah virus; Vaccine design; Immunogenicpeptides; Molecular dynamics

Introduction

Nipah Virus has been declared as a global health problem due to its high mortality in humans and non-availability of its vaccine or other means of prophylaxis (1,2) and was first confirmed to infect humans in Malaysia in 1999 with symptoms of encephalitis (3). Nipah Virus has negative polarity single-stranded RNA (4) with extremely less generation time rendering it more potent to mutate and infect new hosts. It has six genes which encode Nucleocapsid (N), Phosphoprotein (P), Matrix protein (M), Fusion protein (F), Attachment Glycoprotein (G) and RNA polymerase (L) (5). Once internalized, the incubation time for this virus ranges between 4-45 days and the typical symptoms appear only after 4-14 days (2). The virus infection may affect the major organs such as heart, brain, kidneys, lungs, spleen etc. Symptoms of encephalitic syndrome, serious neurological damage and tachycardia have been also reported moreover in Malaysia, more than 50% of patients compromised consciousness whereas in Bangladesh and India over 90% of patients have been reported such severe symptoms (6). This virus is zoonoticand is known to be transmitted from various species of bats (7) and in addition, human-to-human transmission

has also been observed via direct contact with patient secretions (8).

No vaccinations or medications have been approved for use in humans as of yet, despite the fact that several research on the NiV vaccine have been published and moreover, Nipah Virus treatment includes prophylaxis of venous thrombosis, airway patency maintenance and body electrolyte balance maintenance (9).

The current emphasis of this research is the development of the vaccines with enhanced safety profiles and promoting T-cell-mediated immunity. Evaluating T-cell responses has gained popularity in the initial stages of therapeutic trials. The effectiveness of vaccinations in providing long-lasting immunity and generating antibodies that can neutralize a wide range of pathogens is determined by the immunological characteristics of T cells. Research has been conducted to determine the activation state of human B- and T-cell populations following acute Nipah virus infection. Although effector features were seen in activated CD8 T-cells, activated CD4 T-cells were less prominent (10).

Considering the risks of infection, the codon usage of Nipah virus about human hosts was studied to identify the genes with most translation potential to design CD8+ cytotoxic T cell immunogenic peptides for development of vaccine against Nipah virus. The selected peptides were then docked with their MHC-I heavy chain to identify the types of physical interactions followed by molecular dynamics simulations to evaluate the dynamic interactions of the immunogenic peptides with the MHC-I alleles. The selected 11 peptides from the four structural proteins of Nipah virus were predicted to interact with the HLA alleles which cover over 95% to 98% human population worldwide and hence offer a strong case for further development of Nipah virus vaccine.

Materials and Methods

This study was conducted at the School of Biotechnology, Katra, India, in 2022.

Sequence retrieval and host codon usage estimation

Nucleotide and amino acid sequences of all CDSs of the Nipah virus (F, G, C, L, M, N, P, V and W) were retrieved from the NCBI in FASTA format. Net nucleotide distribution of A, T, G and C in the CDSs, especially at the end position of the codons i.e., U3%, G3%, C3% and A3%, net AU% was calculated. Overall distribution of A and U at third position in codons (AU3) as well as G or C at third position (GC3) was calculated. Followed by this, the net G and C present (GC%), G and C at the first two positions of codon (GC12%) and G/C at third position (GC3pos) were calculated. Relative dinucleotide abundance analysis was done to identify the dinucleotide pairs favored by the host organisms. There are 16 dinucleotide combinations for A, T, G and C and the dinucleotide frequency patterns explain the mutational and selection pressures in viruses and Relative synonymous codon usage analysis (RSCU) is the ratio of the observed to the expected value for the amino acid residue. For the overrepresented codons the RSCU value is >1.6, and for the underrepresented codons this value is <0.6. For codons with RSCU values 0.6 and 1.6 are considered as randomly used or unbiased codons. Method developed by Wright was applied to identify the bias in utilization of synonymous codons to identify the effective number of codons (Nc) and its value ranges from 20 to 61, where 20 indicate maximum bias and only one codon per amino acid is used even if other codons are available whereas 61 indicates no bias at all which indicates the equal use of all the codons. Nc > 35 is understood as highly biased codon usage. Relative Codon Deoptimization Index (RCDI) analysis was performed to understand the similarity between the viral genes and host genome to estimate the viral gene's translation potential in host(11).

Selection of antigenic proteins and immunogenic peptides

The antigenicity of the structural proteins was analyzed using the widely known VaxiJen2.0

server (12). This tool can predict antigenic proteins with greater than 80% accuracy where the default cutoff of 0.4 was set for antigenicity calculation and the antigenic structural proteins above the cutoff were used for identification of immunogenic peptides. The physicochemical and structural properties of the antigenic proteins such as molecular weight, amino acid composition, hydrophilicity and hydrophobicity were then computed using the ExpasyProtParam server (13). CD8+ T Cell Epitope (CTL) were identified using the NetCTL 1.2 server at threshold of 0.75 for more conclusive and effective prediction using the Major Histocompatibility Complex Class I (MHC-I) i.e., A1, A2, A3, A24, A26, B7, B8, B27, B39, B44, B58, and B62 (14). For IC50 (half maximal concentration for inhibition), stabilized matrix method (SMM) was employed and IC50 <50 nM was selected as higher affinity in this analysis. The peptides predicted so far were then studied for Antigenicity prediction using the VaxiJenserver(12), Allergenicity prediction using the AllergenFPserver (15) and for Toxicity prediction using the ToxinPred server (16). During viral infection, Helper T Cell epitopes stimulate the release of various cytokines, which further activate other immune cells such as natural killer cells, macrophages, cytotoxic T cells etc., therefore IFN-gamma induction potential of these non-allergenic/non-toxic/antigenic/immunogenic peptides was predicted through IFNepitope web server (17).

MHC-I allele prediction and population coverage

Corresponding MHC/HLA allele that binds to the antigenic peptides were predicted by NetCTL2 server; the IEDBs MHC-I prediction server was used using its recommended settings to identify top of 27 HLA reference set MHC-I alleles(18). The results are reported in percentile rank as a measure of affinity of the peptide with the MHC-I allele, where lower percentile exhibits strong binding. In this study, the percentile rank was fixed at 1.3. Population coverage analysis of vaccine candidate peptides was done using IEDBs server to design a wide spectrum vaccine possessing high efficacy across global ethnicities (19).

MHC-I structure modeling

3D structures of the selected MHC-I alleles were generated through the Swiss-Model and the molecular docking was done using the Autodock-Vina under the default parameters (20–22). The 3D structure of the peptides was modeled using the PepStrMod online server. Criteria for the geometric cutoffs values to identify the interactions between the alpha chain of MHC-I and the antigenic peptides used in this study are as under: H bonds, where a distance of 2.5 Å between the donor and acceptor, donor angle of ≥120º between donor hydrogen acceptor atoms and acceptor angle ≤90º between hydrogen acceptor bonded atoms. Three types of hydrophobic interactions i.e., π-cation, π-π and the other were recorded. For π-cation, aromatic and charged groups exists between 4.5 Å, in π-π interaction, the two aromatic groups should be stacked face to face or face to edge and for other nonspecific hydrophobic interactions, the hydrophobic side chains should be ≤3.6Å distance. Ionic interactions between oppositely charged atoms were considered if their distance was ≤3.7Å. Water bridges were protein peptide interactions mediated by a water molecule. Geometric criteria was D-H..A donor angle ≥110Å, H…A_X angle of ≥90Å.

Molecular dynamics simulations

Desmond molecular dynamics software (academic version) was used to conduct the simulations of Nipah virus peptides bound with MHC-I (23). Overall, 11 MD simulations were conducted and the length of each MD trajectory was 50 nanoseconds. The protein preparation wizard of Schrodinger Maestro was used to optimize the proteins (24). System builder tool was used to place the protein in the center of an orthorhombic box and solvated with water using TIP3P water model. 0.15 M NaCl was added to mimic the physiological salt concentration (23). Overall charge of the system was neutralized using the counter ions of Na^+ or Cl^- ions. The MD simula-

tions were conducted at 300ºK and 1 Bar conditions. The default parameters were used for the preparatory and production phase of MD. One thousand frames were generated from each MD simulation trajectory.

Results

Codon usage analysis

Coding sequences (CDSs) of Nipah virus were studied and observed that A and U nucleotides were in abundance as compared to the G and C nucleotides. The instances of A, U, G and C nucleotides at the third position of the CDSs were observed in the order A>U>G>C as reported in the previous study (25). The Nc-GC3 plot was generated and the dots representing the CDSs below the curve indicating the codon usage were due to the selection pressure (Fig. 1).

Antigenicity prediction

Through the ProtParam tools (13) the characterization of the physicochemical properties (Table 1) of these viral proteins was also done such as the longevity of these proteins was predicted over 30 hours in mammalian cells.

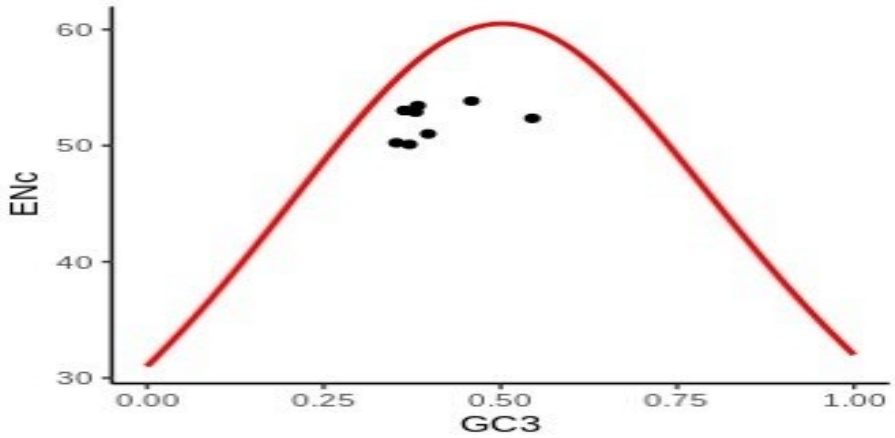

Fig. 1: The Nc-GC3 plot showing the dots representing the CDSs below the curve, indicating the codon usage of the Nipah virus genes were influenced by selection pressure

Table 1: The physicochemical properties and antigenicity score of the antigenic structural proteins of the Nipah Virus

S.No.	*Name*	*Vaxijen antigenicity score*	*Amino acids*	*Mol.Wt (KDa)*	*pI*	*Ext. coff.*	*Half life (Hrs.)*	*Instability Index*	*Aliphatic Index*	*Avg. Hydropathicity*	*-vely charged Amino Acids*	*+vely charged Amino Acids*
1	Nucleoprotein (N)	0.5494	532	58.168	6.06	43890	30	52.33	86.28	-.236	64	62
2	Phosphoprotein (P)	0.5720	709	78.3	4.61	46215	30	48.52	76.29	-.730	128	76
3	Fusion Protein (F)	0.773	546	60.281	5.85	42135	30	38	112.44	0.195	49	46
4	Attachment Glycoprotein (G)	0.5135	602	67.039	8.58	76750	30	34.56	90.95	-.178	53	61

Prediction of CD8+ peptides

From the antigenic proteins, nanomer peptides were identified which could bind to the MHC-I proteins hence using the NetCTL 1.2 web server (14). 126, 105, 133 and 133 potential CD8+ T cell epitopes (CTL 9mer) were predicted using the default threshold of 0.6 for the structural proteins respectively (Table 2).

Table 2: List of top selected CD8+ T Cell Lymphocytes

CD8+ T cell Epitope (CTL)					
Fusion Protein					
Peptide sequence	Antigenicity	Score	Allergenicity	Toxicity	Position
DISSQISSM	Antigen	0.5778	Non-Allergenic	Non-Toxin	455-463
FTDKVDISS	Antigen	0.5729	Non-Allergenic	Non-Toxin	450-458
MLSMIILYV	Antigen	0.446	Non-Allergenic	Non-Toxin	491-499
SITGQIIYV	Antigen	0.5185	Non-Allergenic	Non-Toxin	261-269
LITFISFII	Antigen	0.4535	Non-Allergenic	Non-Toxin	509-517
ITFISFIIV	Antigen	1.3332	Non-Allergenic	Non-Toxin	510-518
KIKSNPLTK	Antigen	0.725	Non-Allergenic	Non-Toxin	47-55
RRVRPTSSG	Antigen	1.6761	Non-Allergenic	Non-Toxin	531-539
Nucleocapsid Protein					
Peptide sequence	Antigenicity	Score	Allergenicity	Toxicity	Position
NRGYLEPMY	Antigen	0.8971	Non-Allergenic	Non-Toxin	351-359
SMGALNINR	Antigen	1.0967	Non-Allergenic	Non-Toxin	344-352
VEETGMAGF	Antigen	0.8613	Non-Allergenic	Non-Toxin	259-267
YLEPMYFRL	Antigen	0.9123	Non-Allergenic	Non-Toxin	354-362
Attachment Glycoprotein					
Peptide sequence	Antigenicity	Score	Allergenicity	Toxicity	Position
LYFPAVGFL	Antigen	0.8143	Non-Allergenic	Non-Toxin	362-370
MGIRPNSHY	Antigen	0.871	Non-Allergenic	Non-Toxin	399-407
MNIMIIQNY	Antigen	0.4923	Non-Allergenic	Non-Toxin	65-73
VMPYGPSGI	Antigen	0.5878	Non-Allergenic	Non-Toxin	348-356
YFPAVGFLV	Antigen	0.8129	Non-Allergenic	Non-Toxin	363-371
Phosphoprotein					
Peptide sequence	Antigenicity	Score	Allergenicity	Toxicity	Position
AQPPYHWSI	Antigen	0.9115	Non-Allergenic	Non-Toxin	360-368
EIQEIANTV	Antigen	0.5284	Non-Allergenic	Non-Toxin	693-701
HLVSMMIMI	Antigen	0.4048	Non-Allergenic	Non-Toxin	570-578
MPKSRGIPI	Antigen	1.4672	Non-Allergenic	Non-Toxin	396-404
QLNASTAVK	Antigen	0.465	Non-Allergenic	Non-Toxin	449-457
RHVRGSPPY	Antigen	0.7335	Non-Allergenic	Non-Toxin	429-437
SEVIVGISP	Antigen	0.9906	Non-Allergenic	Non-Toxin	259-267
VTDVVYHDH	Antigen	0.6573	Non-Allergenic	Non-Toxin	111-119
YPSAGTENV	Antigen	0.5047	Non-Allergenic	Non-Toxin	412-420

Allergenicity, antigenicity, toxicity and immunogenicity prediction

MHC-I binding peptides of structural protein-were identified as non-allergenic using the AllergenFP web server (15). Following, the ToxinPred web server (16) was used to identify 8, 4, 5 and 9 peptides as non-toxic peptides for above proteins (Table 1). Further, 3 peptides for Nucleocapsid and 3 each for Phosphoprotein, Attachment Glycoprotein and 2 peptides for Fusion protein were found to have potential to induce the IFN gamma to ascertain their potential to activate the components of the immune system. All these peptides were predicted as positive inducers of the IFN gamma (Tables 3-6) using the IFNepitop web server (17).

Table 3: Selected CD8+ T Cell Epitopes of Nucleocapsid protein binding alleles of MHC-I

Nucleocapsid Protein		
Peptide	MHC-I allele	Percentile rank (<10)
SMGALNINR	HLA-A*31:01	0.25
	HLA-A*33:01	0.82
	HLA-A*11:01	1.3
VEETGMAGF	HLA-B*44:03	0.33
	HLA-B*44:02	0.35
	HLA-B*40:01	0.44
YLEPMYFRL	HLA-A*02:01	0.05
	HLA-A*02:06	0.11
	HLA-A*02:03	0.29
	HLA-B*08:01	0.78
	HLA-A*23:01	0.82
	HLA-A*24:02	0.85
	HLA-A*32:01	1.1

Table 4: Selected CD8+ T Cell Epitopes of Attachment glycoprotein binding alleles of MHC-I.

Attachment Glycoprotein		
Peptide	MHC-I allele	Percentile rank (<10)
LYFPAVGFL	HLA-A*23:01	0.09
	HLA-A*24:02	0.12
MGIRPNSHY	HLA-A*30:02	0.14
	HLA-B*15:01	0.27
	HLA-B*35:01	0.38
	HLA-B*58:01	0.74
	HLA-B*53:01	0.77
	HLA-A*01:01	0.93
	HLA-A*26:01	0.95
	HLA-B*57:01	1.2
YFPAVGFLV	HLA-A*24:02	0.39
	HLA-A*23:01	0.41

Table 5: Selected CD8+ T Cell Epitopes of Phosphoprotein binding alleles of MHC-I

Phosphoprotein		
Peptide	MHC-I allele	Percentile rank (<10)
EIQEIANTV	HLA-A*68:02	0.02
	HLA-A*26:01	0.45
	HLA-B*51:01	0.75
	HLA-A*02:06	1.3
HLVSMMIMI	HLA-A*02:03	0.99
	HLA-A*02:01	1.1
YPSAGTENV	HLA-B*51:01	0.11
	HLA-B*53:01	0.26
	HLA-B*35:01	0.31
	HLA-B*07:02	0.5
	HLA-A*68:02	0.91

Table 6: Selected CD8+ T Cell Epitopes of Fusion binding alleles of MHC-I

Fusion protein		
Peptide	MHC-I allele	Percentile rank (<10)
DISSQISSM	HLA-A*26:01	0.05
	HLA-B*35:01	0.5
	HLA-A*68:02	0.65
	HLA-B*08:01	0.77
ITFISFIIV	HLA-A*68:02	1.2

Using the IEDB's population coverage tool (19), the peptides of the Nucleocapsid, Phosphoprotein, Attachment Glycoprotein and Fusion protein had 82.2%, 63.2%, 60.58% and 25.22% population coverage (Fig. 2).

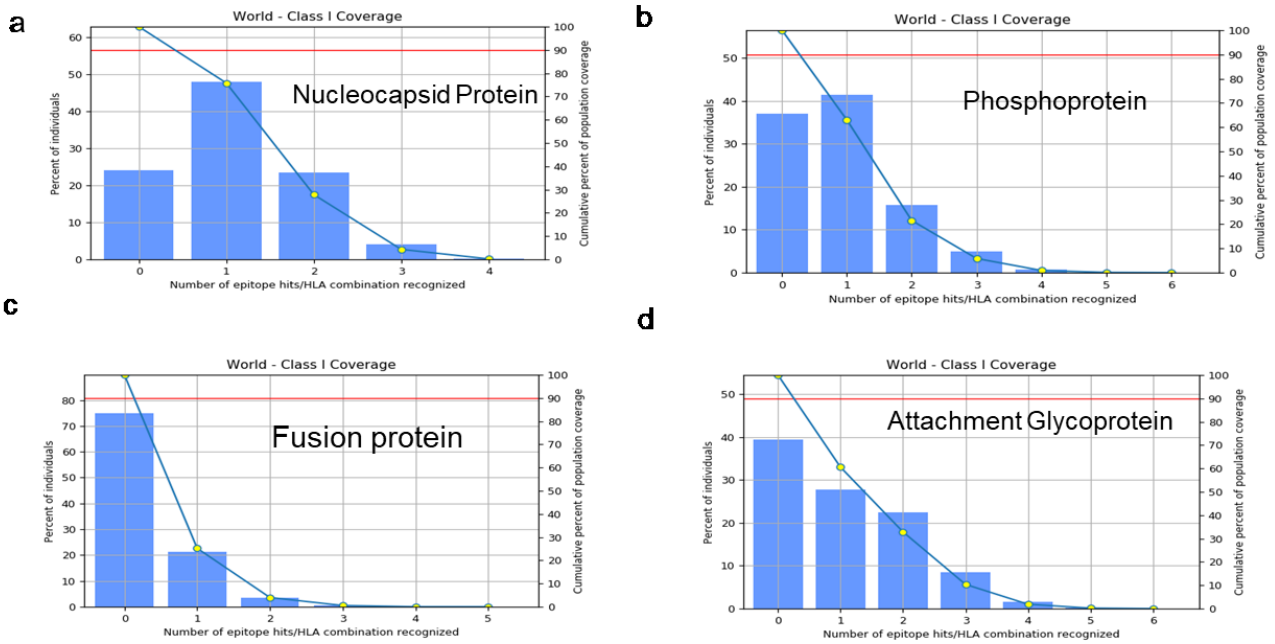

Fig. 2: Bar graphs showing the Human Global Population coverage based on the set of HLA alleles of the IEDB server

Docking and molecular dynamics of selected peptides with MHC-I

Further, using the PyRx and Autodock tool, the docked complexes of the peptides and MHC-I were generated (Fig. 3) and subjected to all atom molecular dynamics simulations; the length of each of these MD simulations was 50 ns. The root mean square deviation plots of the heavy chain of MHC-I alleles were also generated. All the MHC-I heavy chains had converged their RMSD trajectories during the MD simulation (Fig. 4), suggesting the lesser deviation in these receptors upon binding with their peptide molecules. The RMSF plots were also plotted along with their residues interacting with peptides highlighted (Fig. 5). To visualize the interactions of peptide with peptide binding groove, protein-peptide contact diagrams evolved across the length of the MD trajectory were also plotted.

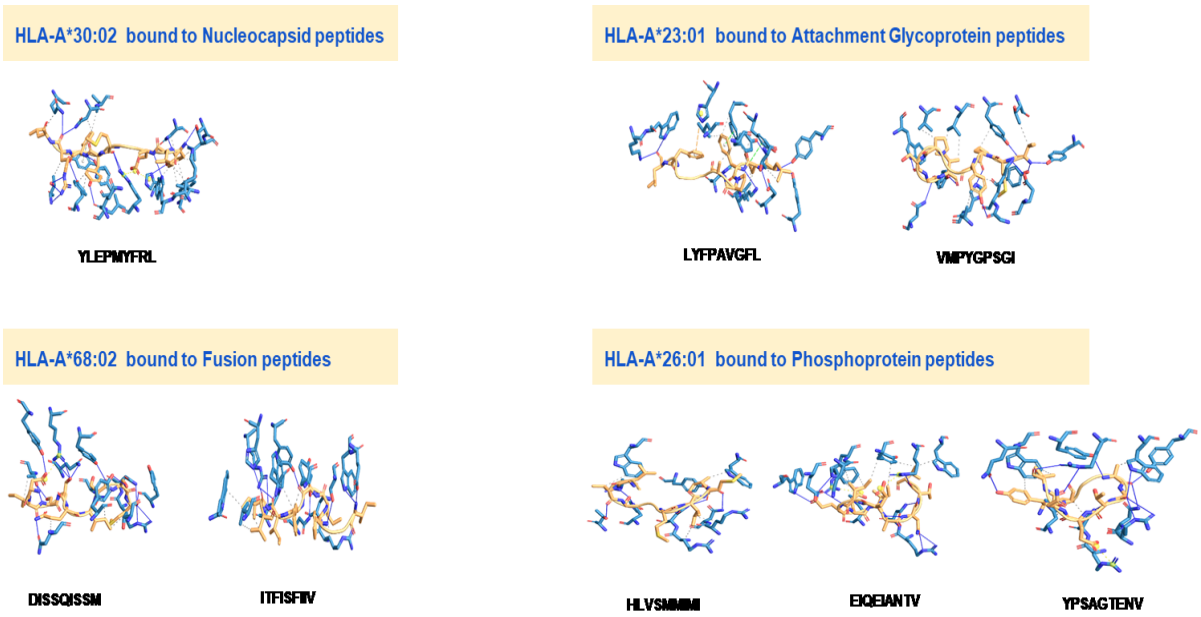

Fig. 3: Binding of HLA alleles with peptides of structural proteins. The HLA-A amino acids are shown in blue while and the immunogenic peptides are shown in yellow stick representation

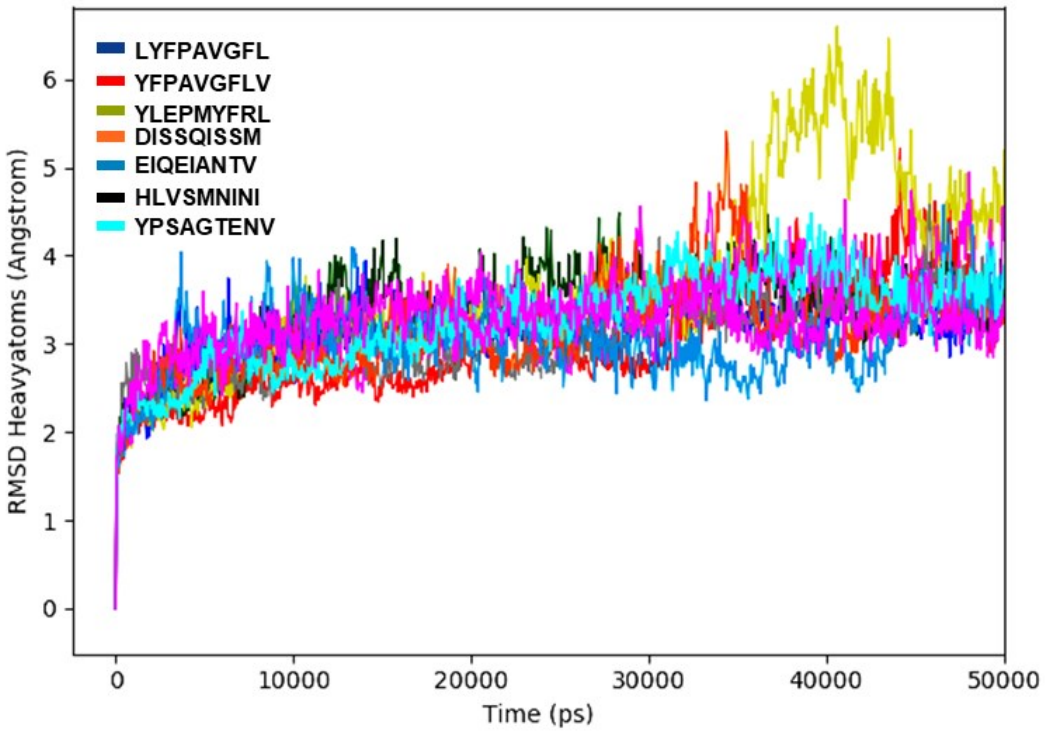

Fig. 4: Nucleocapsid, Attachment Glycoprotein, Fusion peptides, and Phosphoprotein. The convergence of RMSD during the 50 nanoseconds molecular dynamics simulation trajectory hints at structural stability of the MHC-I

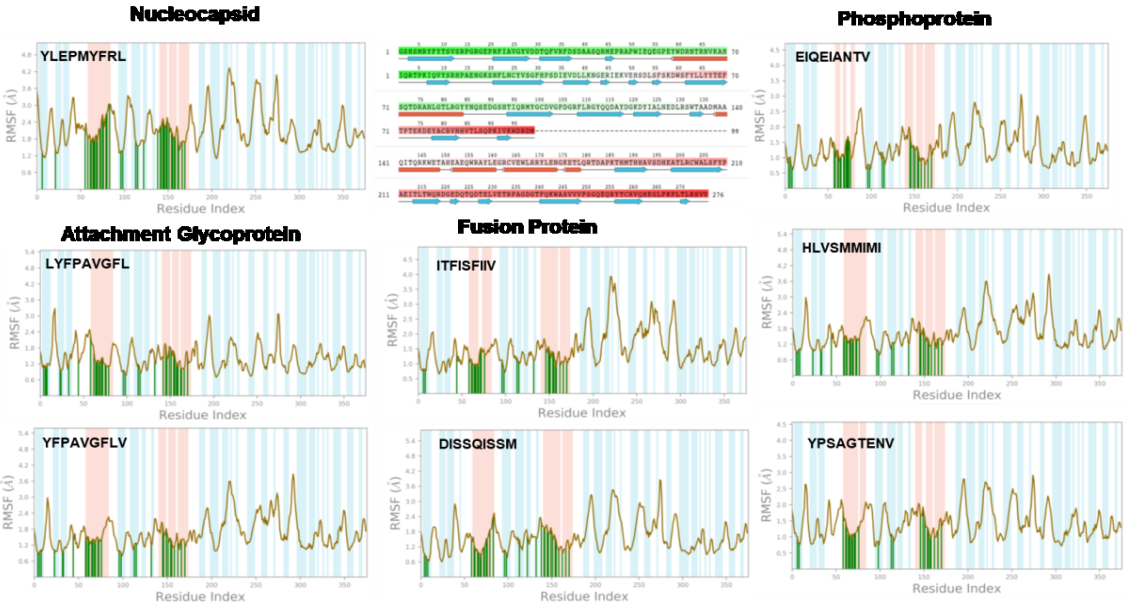

Fig. 5: The root mean square fluctuation (RMSF) plots of MHC-I heavy chain with immunogenic peptides are shown along with the amino acid positions (green) directly involved in interactions with the peptides highlighted. The primary sequence of MHC-I heavy chain is also shown with secondary structure

Nucleocapsid's immunogenic peptides

The top one peptide interacted with the HLA_A_ 30.02 allele. During the 50 ns MD trajectory, five hydrophobic interactions were observed between the YLEPMYFRLI and MHC-I allele's antigen binding pocket residues i.e., Tyr 7, Tyr59, Tyr99, Trp 147 and Tyr159. Nine major water bridges were formed by Tyr 59, Glu63, Asn66, Asn 77, Tyr 80, Tyr 99, Glu114, Arg152, Gln155 and Aln158 amino acids. Tyr 7, Tyr 59, Tyr 99, Trp 147, Leu156 Tyr 159 of the heavy chain formed H-bonds with the YLEPMYFRLI peptide (Fig. 6).

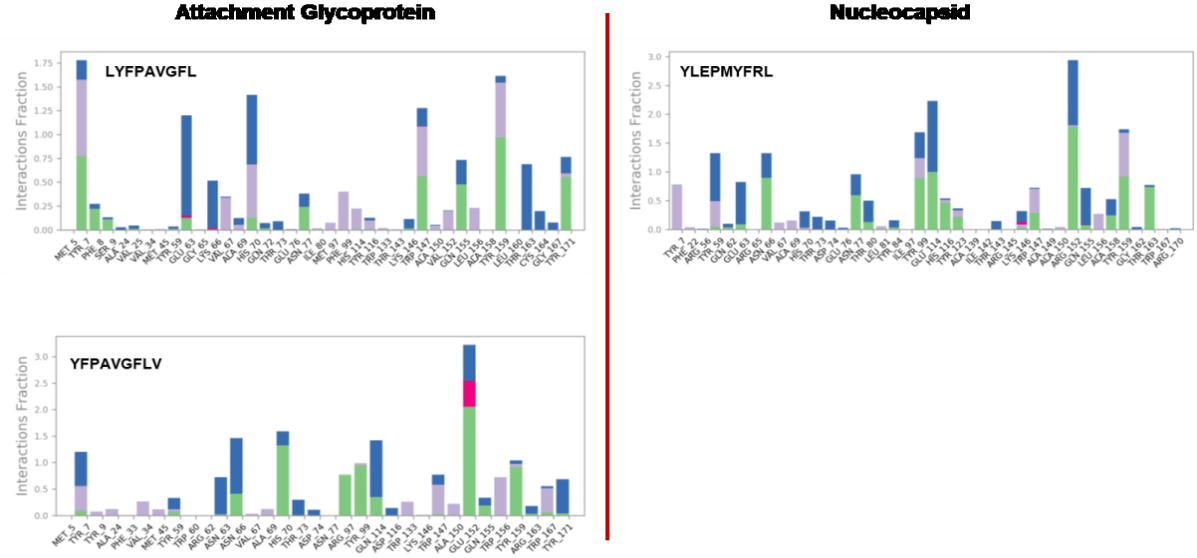

Fig. 6: MHC-I and Nucleocapsid/ Attachment Glycoprotein immunogenic peptide's contact map showing various amino acids of MHC-I interacting with the peptides via H-bond (green), Hydrophobic interactions (purple), water bridge (blue) and ionic interactions (red)

Attachment Glycoprotein's immunogenic peptides

HLA-A-23.011 was predicted to interact with all three peptides of Attachment Glycoprotein. Through MHC-I peptide contacts maps, it was observed that HLA-A-23-01 formed seven H-bonds with the LYFPAVGFL peptide via its Tyr7, Phe8, Asn7, Trp147, Gln155, Tyr159 and Tyr171. Nine hydrophobic interactions were contributed by the Tyr7, Val67, His70, Phe99, His114, Trp147, Val152, Leu156 and Tyr159. Six water bridges were formed by the Glu63, Lys66, His70, Trp147, Gln155, Thr163 and Cys164. MHC-I and YFPAVGFLV peptide's contacts map revealed seven H-bond contacts involving Asn66, His70, Arg97, Tyr99, Gln114, Glu152, and Tyr159 amino acids of MHC-I heavy chain. Six hydrophobic interactions were contributed by the Tyr7, Val34, Trp133, Trp147, Ala150, Trp156 AND Trp167. Single ionic interaction was formed by the Glu152. Seven water bridges were formed by the Tyr7, Asn63, Asn66, Thr73, Gln114, Glu152 and Tryr 171. Around 18 distinct amino were involved in these interactions across the length of the MD trajectory (Fig. 6).

Fusion protein

Two peptides selected for MD simulations with the HLA_A_68_02 allele of MHC-I. HLA_A_68_02 and ITFISFIIV complex's MD was evaluated for the protein-peptide contacts. Seven H bonds were formed by the Asn63, Asn66, Gln70, Tyr99, Gln155, Tyr159 and Trp167 of the MHC-I heavy chain amino acids with the peptide. Three hydrophobic interactions were contributed by the Tyr7, Trp156 and Tyr159. Seven water bridges were formed by the Tyr7, Ala69, Glm70, Thr73, Gln155 and Tyr171. Around 15 amino acids were involved in distinct interaction throughout the MD trajectory timeline (Fig. 7).

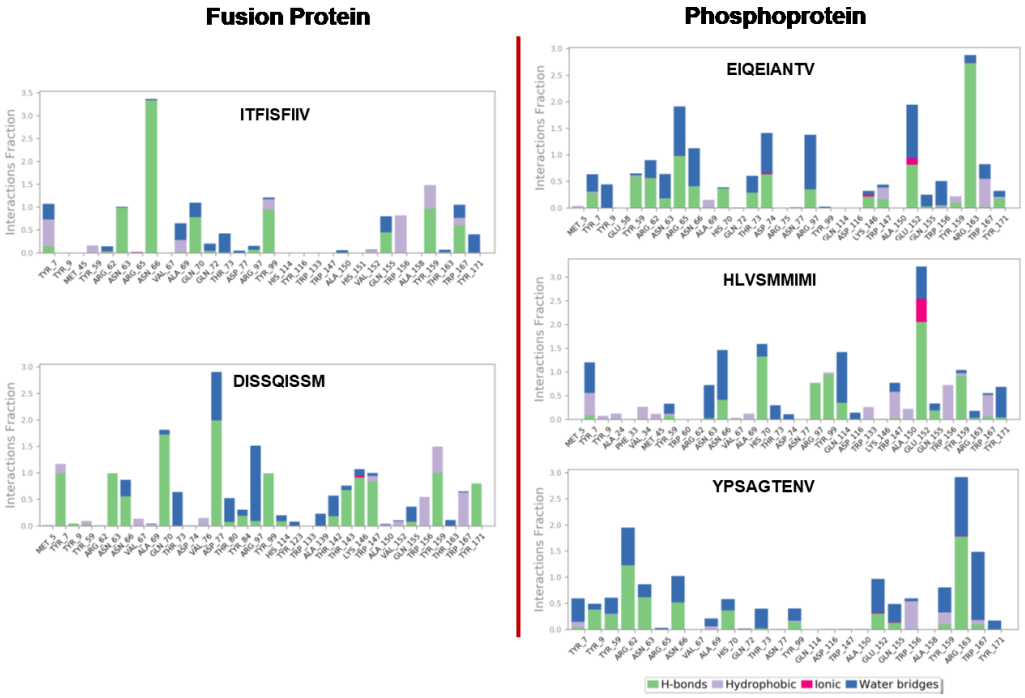

Fig. 7: MHC-I and Fusion peptides/Phosphoprotein immunogenic peptide's contact map showing various amino acids of MHC-I interacting with the peptides via H-bond (green), Hydrophobic interactions (purple), water bridge (blue) and ionic interactions (red)

HLA_A_68_02 and DISSQISSM peptide's contacts map revealed eleven H-bonds formed by Tyr7, Asn63, Asn66, Gln70, Asp77, Tyr99, Thr143, Lys146, Trp147, Tyr159 and Tyr171. Three hydrophobic interactions were formed by the Trp156, Tyr159 and Trp167.Six water bridges were formed by the As66, Thr73, Asp77, Thr80, Arg97, Thr142 and Gln155 of the MHC-I heavy chain amino acids with the immunogenic peptide. Throughout the timeline of the MD trajectory, around 18 distinct amino acids were involved in contact with DISSQISM peptide. This suggests that peptide was accommodated well in the peptide-binding groove (Fig. 7).

Phosphoprotein

Three immunogenic peptides of the phosphoprotein i.e., EIQIEANTV, HLVSMMIMI AND YPSAGTENV were studied through molecular dynamics to visualize their binding to the MHC-I protein. The common HLA allele interacting with these peptides was HLA-A*26:01. HLA-A*26:01 MHC-I complex with EIQIEANTV had nine H-bonds involving Tyr7, Tyr59, Arg62, Arg65, His70, Asp74, Glu152 and Arg163. Hydrophobic interaction was added by Trp167. Ten water bridges were observed where Tyr7, Tyr9, As63, Arg65, Asn66, Thr73, Asp74, Arg97, Glu152 and Trp156 of MHC-I heavy chain were involved. Twenty amino acids of the MHC-I remained involved in consistent interactions with the peptide throughout the MD simulation timeline (Fig. 7). HLA-A*26:01 and HLVSMMIMI complex revealed six H- bonds contributed by the Asn66, His70, Arg97, Tyr99, Glu152 and Tyr159 of MHC-I with HLVSMMIMI. Four hydrophobic interactions were formed by the Tyr7, Trp 147, Trp156 and Trp167 with the peptide. Single ionic interaction was contributed by the Glu152. Tyr7, Asn63, Asn66, Gln114, Glu152 and Tyr171 formed six water bridges with the peptide. Eighteen amino acids of the MHC-I were found to remain in interaction with the peptides throughout the MD simulation highlighting the consistency of the interaction between the two (Fig. 7).

For YPSAGTENV, HLA-A*26:01's six amino acids i.e., Tyr9, Arg62, Asn63, Asn66, His70 and Arg163 contributed in H- bonding with the peptide. Trp156 formed the hydrophobic interaction. Nine ionic interactions were formed through the involvement of the Tyr7, Arg62, Asn66, Thr73, Glu152, Gln155, Tyr159, Arg163 and Trp167 of the MHC-I heavy chain. Consistently, throughout the MD simulation around sixteen amino acids of MHC-I heavy chain remained in contact with the peptide, highlighting the consistent accommodation of peptide in heavy chain's peptide binding groove (Fig. 7). The VEETGMAGF, SMGALNINR and MGIRPNSHY peptides complexed with the MHC-I were not simulated due to computational resource constraints.

Discussion

The Nipah virus (NiV) presents a major public health risk and declared as a global health problem by the WHO and Centre for Disease Control and Prevention (CDC) (1, 2). In the regions such as Australia, Bangladesh, India, Malaysia, and Singapore, where it is linked to high mortality rates due to conditions like acute respiratory distress syndrome and encephalitis (30). The lack of approved vaccines or treatments for human use highlights the urgent need for effective vaccine development. This study offers valuable insights into potential immunogenic peptides that could be crucial for creating a viable vaccine.

One notable aspect of this research is the frequency distribution of viral proteins, which may affect their stability and expression, ultimately influencing the immune response. Understanding the virus's genetic structure is essential for pinpointing conserved regions suitable for vaccine targeting. The identification of potential CD8+ T cell epitopes is especially important, as cytotoxic T cells play a vital role in clearing the virus and preventing severe disease. Notably, the high cumulative population coverage of 94.12% for the selected HLA alleles suggests that the vaccine could be effective across diverse ethnic groups,

making it a strong candidate for global application.

This study underscores the importance of identifying CD8+ T cell epitopes in the development of effective vaccines. By generating robust immune responses, we can enhance protection against the Nipah virus and improve overall vaccination strategies. Focusing on conserved and immunogenic peptides will allow for future vaccine formulations that promote broad and long-lasting immunity. Consequently, the findings from this study represent a significant advancement in the quest for a Nipah virus vaccine. The identified immunogenic peptides could inform future vaccine designs, contributing to public health efforts to control and prevent Nipah virus infections. Ongoing research is crucial for validating these findings and moving toward clinical application.

Conclusion

Bioinformatics workflow was used to identify the immunogenic peptides from the four structural proteins of the Nipah virus. 11 immunogenic, nontoxic, non-allergenic CD8+ antigenic peptides were identified and found to be inducers of interferon gamma. For each of the proteins the resulting peptides were further predicted for their HLA-A alleles and the overall human population coverage against these peptides was predicted using the state-of-the-art tools. Resulting peptides if used in the vaccine will induce an immune response in 92%-98% of the human population across the globe. Further, the docking and molecular dynamics simulation experiments substantiated the strong binding between the MHC-I and peptides.

Journalism Ethics considerations

Ethical issues (Including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, redundancy, etc.) have been completely observed by the authors.

Conflict of interest

The authors declare that there is no conflict of interest.

References

1. Anderson DE, Islam A, Crameri G, et al (2019). Isolation and full-genome characterization of Nipah viruses from bats, Bangladesh. *Emerg Infect Dis*, 25(1):166-170.
2. Ochani RK, Batra S, Shaikh A, Asad A (2019). Nipah virus-the rising epidemic: a review. *Infez Med*, 27(2):117–27.
3. Chua KB, Goh KJ, Wong KT, et al (1999). Fatal encephalitis due to Nipah virus among pig-farmers in Malaysia. *Lancet*,354(9186):1257–9.
4. Carrasco-Hernandez R, Jácome R, López Vidal Y, et al (2017). Are RNA viruses candidate agents for the next global pandemic? A review. *ILAR J*,58(3):343-58.
5. Martinez-Gil L, Vera-Velasco NM, Mingarro I (2017). Exploring the human-Nipah virus protein-protein interactome. *J Virol*, 91(23):e01461-17.
6. Sharma V, Kaushik S, Kumar R, et al (2019). Emerging trends of Nipah virus: A review. *Rev Med Virol*,29(1):e2010.
7. Sudeep AB, Yadav PD, Gokhale MD, et al (2021). Detection of Nipah virus in Pteropusmedius in 2019 outbreak from Ernakulam district, Kerala, India. *BMC Infect Dis,*21(1):162.
8. Gurley ES, Montgomery JM, Hossain MJ, et al (2007). Person-to-person transmission of Nipah virus in a Bangladeshi community. *Emerg Infect Dis,* 13(7):1031-7.
9. Skowron K, Bauza-Kaszewska J, Grudlewska-Buda K, et al (2022). Nipah virus–Another threat from the world of zoonotic viruses. *Front Microbiol*, 12:811157.
10. Arunkumar G, Devadiga S, McElroy AK, et al (2019). Adaptive Immune Responses in Humans During Nipah Virus Acute and Convalescent Phases of Infection. *Clin Infect Dis*, 69(10):1752–6.
11. Anwar AM, Soudy M, Mohamed R (2019).vhcub: Virus-host codon usage co-adaptation analysis. *F1000Res,*8:2137.

12. Doytchinova IA, Flower DR (2007).VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinformatics*, 8:4.
13. Gasteiger E, Hoogland C, Gattiker A, et al (2005). Protein identification and analysis tools on the ExPASy server. In: Walker, J.M. (eds) *The Proteomics Protocols Handbook*. Springer Protocols Handbooks. Humana Press.
14. Larsen M V, Lundegaard C, Lamberth K, et al (2007). Large-scale validation of methods for cytotoxic T-lymphocyte epitope prediction. *BMC Bioinformatics*, 8:424.
15. Dimitrov I, Naneva L, Doytchinova I, et al (2014). AllergenFP: allergenicity prediction by descriptor fingerprints. *Bioinformatics*, 30(6):846–51.
16. Gupta S, Kapoor P, Chaudhary K, et al (2013).In silico approach for predicting toxicity of peptides and proteins. *PLoS One*, 8(9):e73957.
17. Dhanda SK, Vir P, Raghava GPS (2013). Designing of interferon-gamma inducing MHC class-II binders. *Biol Direct*, 8:30.
18. Lundegaard C, Lund O, Nielsen M (2008). Accurate approximation method for prediction of class I MHC affinities for peptides of length 8, 10 and 11 using prediction tools trained on 9mers. *Bioinformatics*, 24(11):1397–8.
19. Bui HH, Sidney J, Dinh K, et al (2006). Predicting population coverage of T-cell epitope-based diagnostics and vaccines. *BMC Bioinformatics*,7:153.
20. Waterhouse A, Bertoni M, Bienert S, et al (2018). SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res*, 46(W1):W296–303.
21. Bienert S, Waterhouse A, De Beer TAP, et al (2017). The SWISS-MODEL Repository—new features and functionality. *Nucleic Acids Res*, 45(D1):D313–D319.
22. Trott O, Olson AJ (2010). AutoDockVina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem*,31(2):455–61.
23. Bowers KJ, Chow E, Xu H, et al (2006). *Proceedings of the 2006 ACM/IEEE Conference on Supercomputing*. ACM New York, NY, USA.
24. MadhaviSastry G, Adzhigirey M, Day T, et al (2013). Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J Comput Aided Mol Des*, 27(3):221–34.
25. Khandia R, Singhal S, Kumar U, et al (2019). Analysis of Nipah virus codon usage and adaptation to hosts. *Front Microbiol*, 10:886.
26. Andreatta M, Nielsen M (2016). Gapped sequence alignment using artificial neural networks: application to the MHC class I system. *Bioinformatics*, 32(4):511–7.
27. Wang L, Harcourt BH, Yu M, et al (2001). Molecular biology of Hendra and Nipah viruses, *Microbes Infect*, 3 (4):279–287.
28. Chadha MS, Comer JA, Lowe L, et al (2006). Nipah virus associated encephalitis outbreak, Siliguri, India. *Emerg Infect Dis*, 12 (2): 235–240.
29. Gurley ES, Montgomery JM, Hossain MJ, et al (2007). Person-to person transmission of Nipah virus in a Bangladeshi community. *Emerg Infect Dis*, 13 (7): 1031–1037.
30. Aditi, Shariff M (2019). Nipah virus infection: A review. *Epidemiol Infect*, 147: e95.