In Silico Pharmacology (2024) 12:69
https://doi.org/10.1007/s40203-024-00246-9

ORIGINAL RESEARCH

Revolutionizing Nipah virus vaccinology: insights into subunit vaccine development strategies and immunological advances

Tapas Das[1] · Sutapa Datta[1] · Arnab Sen[1,2,3]

Received: 16 January 2024 / Accepted: 16 July 2024
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2024

Abstract

The Nipah virus (NiV), a zoonotic virus in the Henipavirus genus of the Paramyxoviridae family, emerged in Malaysia in 1998 and later spread globally. Diseased patients may have a 40- 70% chance of fatality depending on the severity and early medication. The recent outbreak of NiV was reported in Kerala (India) by a new strain of MCL-19-H-1134 isolate. Currently, no vaccines are available, highlighting the critical need for a conclusive remedy. Our study aims to develop a subunit vaccine against the NiV by analyzing its proteome. NiV genome and proteome sequences were obtained from the NCBI database. A phylogenetic tree was constructed based on genome alignment. T-cell, helper T-cell, and B-cell epitopes were predicted from the protein sequences using NetCTL-1.2, NetMHCIIPan-4.1, and IEDB servers, respectively. High-affinity epitopes for human receptors were selected to construct a multi-epitope vaccine (MEV). These epitopes' antigenicity, toxicity, and allergenicity were evaluated using VaxiJen, AllergenFP-v.1.0, and AllergenFP algorithms. Molecular interactions with specific receptors were analyzed using PyRx and ClusPro. Amino acid interactions were visualized and analyzed using PyMOL and LigPlot. Immuno-simulation was conducted using C-ImmSim to assess the immune response elicited by the MEV. Finally, the vaccine cDNA was inserted into the pET28a(+) expression vector using SnapGene tool for in silico cloning in an E. coli host. The potential for an imminent outbreak cannot be overlooked. A subunit vaccine is more cost-effective and time-efficient. With additional in vitro and in vivo validation, this vaccine could become a superior preventive measure against NiV disease.

Keywords Nipah virus · Subunit vaccine · Toll-like receptor 3 · Multi epitope · In-silico cloning · Immune simulation

Introduction

Nowadays, numerous infectious diseases have become a significant threat worldwide. Among these Emerging infectious diseases (EIDs), approx. 75% are considered to be zoonotic (transmitted from animals to humans) (Gebreyes et al. 2014). Like the NiV, other viruses like HIV, novel coronavirus, and influenza virus are all examples of zoonotic viruses that cause devastating effects on the human population and cause lethal diseases (Bardhan et al. 2023; Sartorius et al. 2021). Nipha virus is one of them that emerged in Malaysia in 1998 (Hsu 2004). Later NiV outbreaks have been reported worldwide, including in Bangladesh (2001–2005), followed by recently in Kerala, India, before/on September 2023 with the novel strain MCL-19-H-1134 isolate (Chong et al. 2001; Chapa and Garza 2023). It is a member of Paramyxoviridae family and falls under the Henipavirus genus, alongside another prototype virus within the genus known as the Hendra virus (Chua et al. 2000). NiV is an RNA genome-containing virus that possess a total of six genes, namely, the P gene, M gene, N gene, F gene, G gene, and L gene, which codes for nine consecutive proteins (Khandia et al. 2019). The G gene encodes the glycoprotein, which guides the virus in recognizing and attaching to host receptors. Similarly, the F gene codes for the fusion protein, facilitating the fusion of the viral membrane with the host membrane. The N gene encodes nucleoprotein, which remains bound to the RNA, aiding in viral genome

✉ Arnab Sen
arnab.nbu@gmail.com

1 Department of Botany, University of North Bengal, Siliguri, India

2 Bioinformatics Facility Centre, University of North Bengal, Siliguri, India

3 Biswa Bangla Genome Centre, University of North Bengal, Siliguri, India

Published online: 27 July 2024

replication and packaging. The RNA-dependent RNA polymerase enzyme, necessary for replicating the RNA genome, is translated from the L gene. The matrix protein required for virus budding and assembly is encoded by the M gene. Additionally, the P gene is exceptional as it codes not only for the phosphoprotein but also for three additional proteins, C, V, and W, each with a distinct immune modulatory role (Satterfield et al. 2015). Individual epitopes can be screened from non-allergen proteins with further molecular interaction with different MHC complexes (class I and II), Toll-like receptor (TLR-3) protein, and TAP cavity.

The need for a NiV vaccine arises from the severe and often fatal nature of infections. NiV can cause a range of illnesses, including asymptomatic infection, acute respiratory illness, and fatal encephalitis (Banerjee et al. 2019). The high mortality rate associated with NiV infections, as well as the potential for person-to-person transmission, underscores the importance of developing a vaccine to prevent outbreaks and protect populations at risk (Aqsha et al. 2023). Subunit epitope-based vaccines can be a good alternative to traditional vaccines regarding precision and safety (De Groot et al. 2009). It is considered safe for individuals with weakened immune systems or other health concerns, making it suitable for high-risk populations. Subunit epitope-based vaccines can often be developed more rapidly than traditional vaccines. This is important in emerging infectious diseases or outbreaks, where a timely response is critical. In this study, we have predicted all the possible T-cell, helper T-cell, and B-cell epitopes of the W protein of the NiV and shortlisted the major interacting epitopes based on good docking score. The selected epitopes were aligned with the necessary linker sequences to construct a MEV. Adjuvants were incorporated to enhance the immune response. Additionally, the molecular interaction between the MEV and the Human Major Histocompatibility Complex (MHC) was analyzed. The MEVs were administered into the human system using in silico simulation, and the subsequent human immune responses were observed. Finally, complementary DNA (cDNA) was synthesized from the vaccine sequence and inserted into an expression vector for in silico cloning verification.

Material methods

Collection of viral genomes

The entire genome of Henipavirus nipahense (NiV) strains, as well as some other viruses within the Paramyxoviridae family, was sourced from the NCBI database (https://www.ncbi.nlm.nih.gov/). In total, there were 331 genomes documented in the NCBI database prior to 2023, consisting of 62 complete sequences and 279 partial sequences. We specifically extracted viral strains with complete genome sequences. During the 2023 outbreak, a new strain (MCL-19-H-1134 isolate) emerged in India (Kerala) and was reported. This novel isolate, MCL-19-H-1134, was retrieved from the database and subjected to further analysis. While the complete genome sequencing of this novel isolate has not been performed yet, but its partial genome length closely resembles that of NiV with a complete sequence. Viruses like Hendra, Cedar, Ghana, and Mojiang belonging to the NiV family were downloaded from the NCBI database with their complete proteome sequence for the subsequent evolutionary analysis.

Collection of viral proteomes

All the protein-coding genes of NiV were identified and their complete protein sequences were obtained. In this research, a total of nine proteins from the NiV strain were selected for the vaccine epitope. These proteins consist of the Nucleocapsid protein (WKR82401.1), Phosphoprotein (WKR82402.1), V protein (WKR82403.1), W protein (WKR82404.1), C protein (WKR82405.1), Matrix protein (WKR82406.1), Fusion protein (WKR82407.1), attachment glycoprotein (WKR82408.1), and polymerase (WKR82409.1). These nine proteins from the NiV strain were retrieved from the NCBI database. The three-dimensional structures of all nine proteins were obtained from the Protein Data Bank (PDB) server (https://www.rcsb.org/). In cases where the 3D structures of these proteins were not readily available, homology models were created using Swiss-model.

Phylogenetic analysis

The complete genome of NiV isolates from different topological regions of India, Bangladesh, Thailand, Malaysia, and USA have been retrieved from NCBI. For this study, the genome of all NiV isolates is aligned to ensure that homologous regions (regions with common ancestry) are matched up. Genome alignment was done by MAFFT web server (https://www.ebi.ac.uk/Tools/msa/mafft/) from EMBL-EBI (Madeira et al. 2022). Aligned genome data are used to build phylogenetic tree by character-based method. Character-based methods like Maximum Likelihood or Maximum Parsimony estimate the most likely tree from given data. Randomized Axelerated Maximum Likelihood- RAxML (Kozlov et al. 2019) server (https://raxml-ng.vital-it.ch/#/) is used for tree building. Later visualization of the tree has been done by iTOL (Interactive Tree of Life) tool (https://itol.embl.de/upload.cgi) (Letunic and Bork 2021).

Allergenicity of viral proteins

It is imperative in the present study to identify the allergenic protein from the set of data. World Health Organization (WHO) recently focused on the correct prediction of allergenicity in natural as well as modified proteins. In our study we only consider the non-allergenic proteins from MCL-19-H-1134 isolate 2023.

An auto-cross covariance (ACC) is a protein sequence mining method that is used for uniformation of protein length (Wold et al. 1993). Each amino acid present in a protein sequence is represented by five important descriptors (represented by E) (Venkatarajan and Braun 2001), later the sequence is transformed into equal vectors by ACC transformation. Five E descriptors defined by Venkatarajan and Braun are the hydrophobicity of amino acids, their helix-forming propensity, correlates with the relative abundance of amino acids, their size and fifth dominated by the β-strand forming propensity. AllergenFp v.1.0 is a Bioinformatics tool for allergenicity prediction that fulfils all the above-mentioned criteria and predicts the allergenicity of the protein data set (https://ddg-pharmfac.net/AllergenFP/). Also, this server predicts the identity level between two different sets by calculating Tanimoto coefficients (Godden et al. 2000). The protein which found allergenic has been eliminated for further study. The overall methodology is given in the flowchart Fig. 1.

Prediction of cytotoxic T-cell (CTL) epitope

Accurate forecasts of Cytotoxic T lymphocyte (CTL) epitopes play a crucial role in the rational design of vaccines. Most importantly, they can reduce the need for extensive experimental work in epitope identification. The MHC-I plays a crucial role in activating Cytotoxic T-cells by serving as antigen presenters. Cytotoxic T cell is a vital

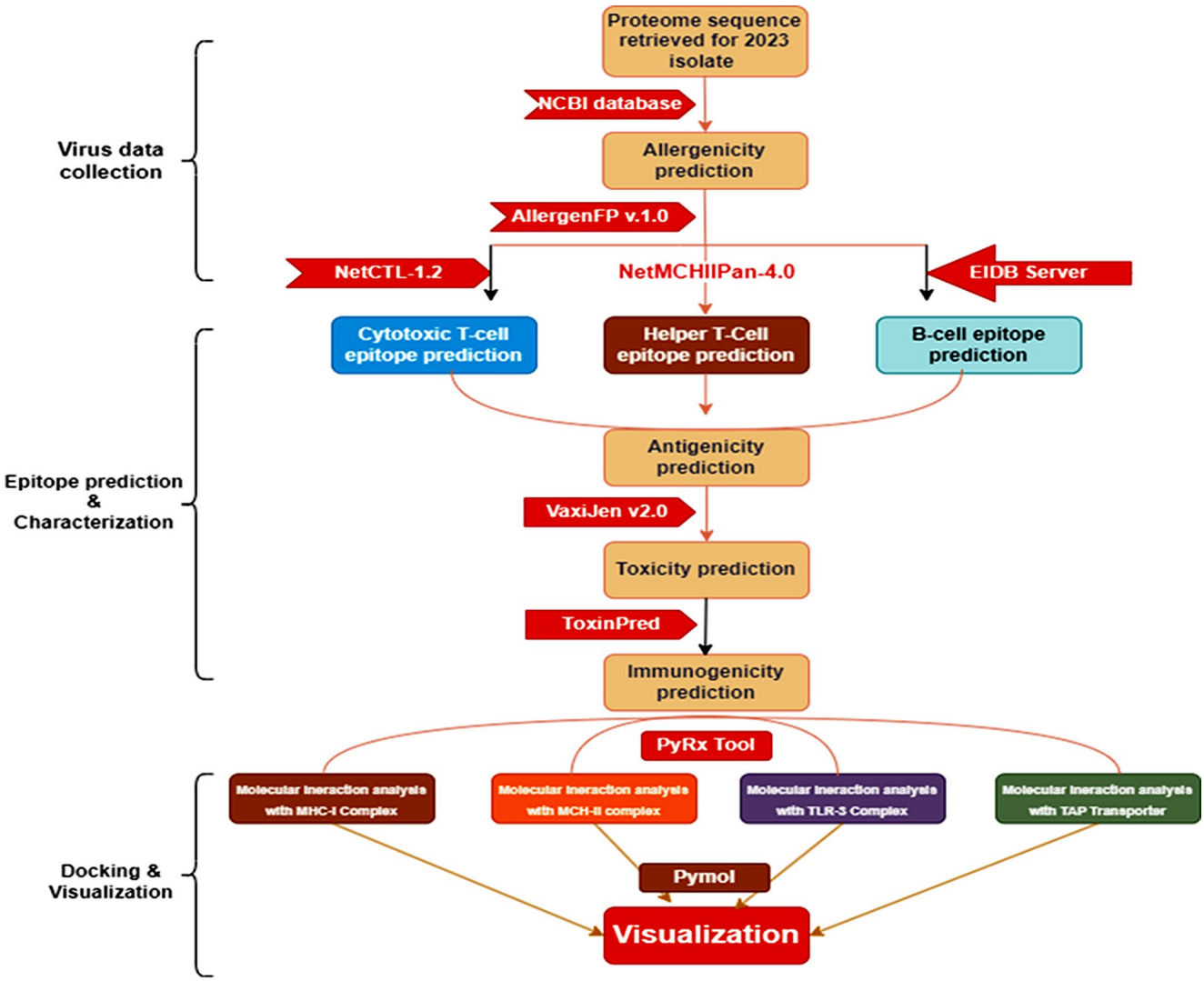

Fig. 1 A flow chart showing the methodology applied to identify the peptide epitope and its interaction with specific receptors

component of the human adaptive immune system they eradicate pathogens from the infected host cell. In the post-infection phase initially, the antigen is engulfed by the proteasome followed by processing in the cytoplasm. After processing these antigens are transported by Transporter Associated with Antigen Processing (TAP). Molecular interaction of TAP transporter with antigenic epitope was also done in this study. For the estimation of CTL, NetCTL-1.2 (Larsen et al. 2007) bioinformatics server was used (https://services.healthtech.dtu.dk/services/NetCTL-1.2/). NetCTL is an online tool specifically created for foreseeing human CTL epitopes within any given protein. It accomplishes this by combining predictions for proteasomal cleavage, TAP transport efficiency, and MHC class I affinity. This method encompasses forecasts for peptide MHC class I binding, proteasomal C-terminal cleavage, and TAP transport efficiency. The server offers predictions for CTL epitopes that are limited to 12 MHC class I supertypes. MHC class I binding and proteasomal cleavage are executed using artificial neural networks, while TAP transport efficiency is anticipated through a weight matrix. For this study weight on C terminal cleavage was taken 0.15, the weight on TAP transport efficiency was taken 0.05, and the threshold for epitope identification was taken 0.75. Furthermore, the epitopes forming strong bonds with the MHC-I supertype were taken under consideration.

Prediction of helper T cell (HTL) epitope

Indeed, the Major Histocompatibility Complex class II (MHC II) plays a pivotal role in the immune system. MHC II molecules are responsible for processing and presenting epitopes from foreign pathogens to Helper T Lymphocytes (HTL). When helper T cells are activated through this interaction, they initiate the generation of adaptive immunity by producing various cytokines. Activated helper T cells are versatile and have a significant impact on the immune response. They assist in activating the humoral immune system by helping B cells with processes like antibody class-switching and affinity maturation. Additionally, they contribute to cell-mediated immunity by activating T-cytotoxic cells and macrophages, enhancing the body's ability to combat infections. One of their most critical roles is in generating memory immune cells. This memory formation is essential for effective vaccines, as it enables the immune system to remember and respond more swiftly and effectively to previously encountered pathogens. Helper T cells are indeed integral in vaccine development and the body's defence against various diseases. For HTL epitope prediction NetMHCIIpan-1.4 tool is used (https://services.healthtech.dtu.dk/services/NetMHCIIpan-4.1/) (Reynisson et al. 2020). NetMHCIIpan-4.1 server is a tool that leverages Artificial Neural Networks (ANNs) to predict peptide binding to various MHC class II molecules. With its extensive training dataset covering HLA-DR, HLA-DQ, and HLA-DP isotypes, it provides valuable insights into the likelihood of a peptide being naturally presented by a specific MHC II receptor. Threshold for strong binder (% Rank) was taken 1 and Threshold for weak binder (% Rank) was taken 5. It allows for more accurate predictions of peptide-MHC II interactions, which are crucial for understanding immune responses and designing effective therapies. In this study epitopes that bind strongly with MHCII are taken into consideration for further study.

B cell epitope prediction

B cell is responsible for the production of antibodies against any pathogen, for that it needs to be activated by a specific epitope. B cell epitope prediction based on the sequence has been done for antigenic proteins of NiV. Antibody epitope prediction has been done from IEDB server by Bepipred Linear Epitope Prediction 2.0 method (http://tools.iedb.org/bcell/) (Jespersen et al. 2017). The tool employs both the propensity scale method and the physiochemical properties of the antigenic sequence for screening potential epitopes. This approach can enhance the accuracy of epitope prediction and aid in various applications, such as vaccine development or immunological studies.

Population coverage analysis of the epitope

T lymphocytes recognize pathogen-derived epitopes through specific MHC molecules. The diversity of these MHC molecules, known as HLA in humans, is substantial, with over a thousand allelic variants identified (Janeway et al. 2001). These variants vary frequently across ethnicities (Al Naqbi et al. 2021). Consequently, when designing T-cell epitope-based diagnostics or vaccines, choosing epitopes that can bind to various HLA alleles is crucial. This approach ensures a broader patient population coverage and is particularly important when dealing with diseases affecting diverse ethnic groups. Population coverage of all the selected epitopes is done by IEDB server (http://tools.iedb.org/population/) (Bui et al. 2006). Epitopes having large population coverage are taken for further analysis.

Characterization of epitope

Antigenicity prediction of epitope

The sequence that is to be used as an epitope should be antigenic to create an immune response in the human

body. All nine proteins of NiV is run through VaxiJen v2.0 tool (https://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html) to predict antigenicity (Doytchinova and Flower 2007). A threshold antigen score of 0.7 is taken for the prediction. Proteins below this score are considered non antigenic and are not incorporated in later analysis.

Toxicity prediction of epitope

Epitopes that are analysed further has to be non-toxic. Toxicity prediction of the epitope is done by the ToxinPred tool (https://webs.iiitd.edu.in/raghava/toxinpred/protein.php). The tool predicts toxicity along with other important physio-chemical properties like hydrophobicity, hydrophilicity, charge, molecular weight, and hydropathicity. The toxicity prediction analysis was done by the "SVM (Swiss-Prot) based" (support vector machine) method.

Immunogenicity prediction

Quantification of immunogenicity is important to determine the level of immune response elicited by an epitope/MHC. Immunogenicity of all the selected epitopes of NiV has been done by the IEDB server (http://tools.iedb.org/immunogenicity/) (Calis et al. 2013).

Molecular docking analysis of CTL and HTL with HLA allele

The interaction between proteins and peptides plays a vital role in cellular functions. Predicting these interactions is essential for comprehending the molecular mechanisms underlying biological processes and facilitating the development of peptide-based vaccine. In this context, the study utilized the PyRx tool (Dallakyan and Olson 2015) to explore the interactions between HLA proteins and identified peptides, which serve as epitopes for T cells. This tool assesses potential interactions between a peptide and its target receptor, constructing optimal models. For the molecular docking analysis, all receptor proteins (MHC-I/MHC-II) were sought in the Protein Data Bank (PDB) server, and their 3D structures were retrieved. Only peptides demonstrating high binding affinity to their respective MCH complexes were included in the study. The HDOCK server (http://hdock.phys.hust.edu.cn/) and GALAXYWEB tool were employed to reinforce the results. The outcomes were visualized using Pymol (Yuan et al. 2017), followed by Ligplot to illustrate interactive amino acids with the receptor molecule.

Molecular interaction analysis of epitopes with TLR-3

The molecular interaction analysis of the selected epitopes with TLR3 was conducted through molecular docking. Initially, HDOCK was employed, followed by the PyRx tool (Dallakyan and Olson 2015). The 3D structure of human TLR3 was obtained from the PDB databank for the molecular docking process. Subsequently, residue interactions across the interface of the chosen epitope in complex with TLR3 were scrutinized using Pymol and visually represented through the Ligplot server.

Analysis of molecular interactions between epitopes and the TAP transporter

The molecular docking analysis of the selected epitopes with the TAP transporter cavity was conducted through molecular docking, utilizing the PyRx tool (Dallakyan and Olson 2015). The cryo-EM structure of the TAP transporter downloaded from PDB (PDB ID: 5u1d) served as the structural model, antigens are removed from the TAP cavity of the original structure and epitope-TAP interactions were studied. The TAP transporter plays a crucial role in presenting CTL epitopes. After the immuno-proteasomal processing of xenobiotic proteins, the fragmented peptides are transported to the endoplasmic reticulum (ER) through the TAP transporter (Abele and Tampe 2004).

Multi-epitope vaccine designing and characterization

Multi-epitope vaccine formulation

For designing a MEV of the NiV selected two CTLs, two HTLs, and two linear B-cell epitopes are aligned in a single line. The correct adjuvants and adaptable interconnected linker sequences are used to formulate the MEV. Due to the lack of immunological properties of epitopes, adjuvants are necessary to trigger the immune response. From N terminus β-defensin, T helper adjuvant (TpD) and *Escherichia coli (E.coli)* FimH adjuvant are aligned. TpD is a universal adjuvant for TCD4 cells, protects mucosal membranes, and stimulates the generation of neutralizing antibodies as well as potentially exhibiting superior performance compared to "PADRE," a peptide that binds to multiple HLA-DR molecules with broad specificity (Alexander et al. 1994; Chawla et al. 2023). FimH adjuvant promotes IFN-γ and TNF-α production; also, it is effective in promoting mucosal immunity (Zhang et al. 2022). To construct a flexible protein configuration for

the vaccine, EAAAK linkers link each adjuvant. CTLs are linked by AAY linkers, HTLs are linked by GGGGS linkers, and KK linkers link B cell epitopes.

Properties of the constructed vaccine candidate

The ProtPharam tool (https://web.expasy.org/protparam/) (Gasteiger et al. 2005) is used to analyze the physiochemical properties of the vaccine. Properties like Molecular weight, number of amino acids, Total number of atoms, Extinction coefficient, Theoretical pI, Aliphatic index (AI), and Instability Index (II) are calculated using the tool. Besides that, a vaccine should be non-allergenic and have antigenic properties. VaxiJen v2.0 (Doytchinova and Flower 2007) and AllergenFP server are used to evaluate the antigenicity and allergenicity of the formulated vaccine candidate.

In-silico Immune simulation of MEV

Immune stimulation of the vaccine is done on the C-ImmSim server (https://kraken.iac.rm.cnr.it/C-IMMSIM/index.php); this server works based on a position-specific scoring matrix (Rapin et al. 2010). The C-ImmSim analyzes the effect of any antigen on several immune cell lineages of the human immune system. Typically, it is advised to space consecutive vaccine doses at least four weeks apart, although in specific cases, a longer interval might be appropriate (Robinson et al. 2018). Therefore, in this study, three antigen injections are given at a time interval of 4 weeks. Simulation is done up to 300 time steps (1 steps = 8 h in real life) injecting antigen at 1, 90, and 180 time steps.

Structure prediction, refinement, and stability of MEV candidate

The term tertiary structure pertains to the overall spatial arrangement of a protein's polypeptide chains in three dimensions. In our research, we utilized the Scratch protein predictor server (https://scratch.proteomics.ics.uci.edu/) to forecast the tertiary structure of a vaccine candidate. To improve the accuracy of our model, we subjected the generated 3D structure of the final vaccine, obtained from the Scratch protein predictor tool, to the GalaxyRefine 2 server (https://galaxy.seoklab.org/cgi-bin/submit.cgi?type=REFINE) for protein structure refinement (Heo et al. 2013). This server operates on data from molecular dynamics simulations following reconstructions of active side chains. The information provided by the server includes MolProbity for depicting crystallographic resolution, global distance test-high accuracy (GDT-HA), root-mean-square deviation (RMSD), and Ramachandran's favored score. RMSD quantifies the distance between atoms, with lower scores indicating more excellent stability; typically, an RMSD falling from 0 to 1.2 is considered acceptable (Heo et al. 2013). Additionally, to assess the stability of the structure, the refined PDB files were analyzed using the PDBsum server (https://www.ebi.ac.uk/thornton-srv/databases/pdbsum/). This tool identifies the most favored regions via Ramachandran plots and examines secondary structures, including motifs such as beta turns, gamma turns, beta-hairpin, and disulfide bond formations.

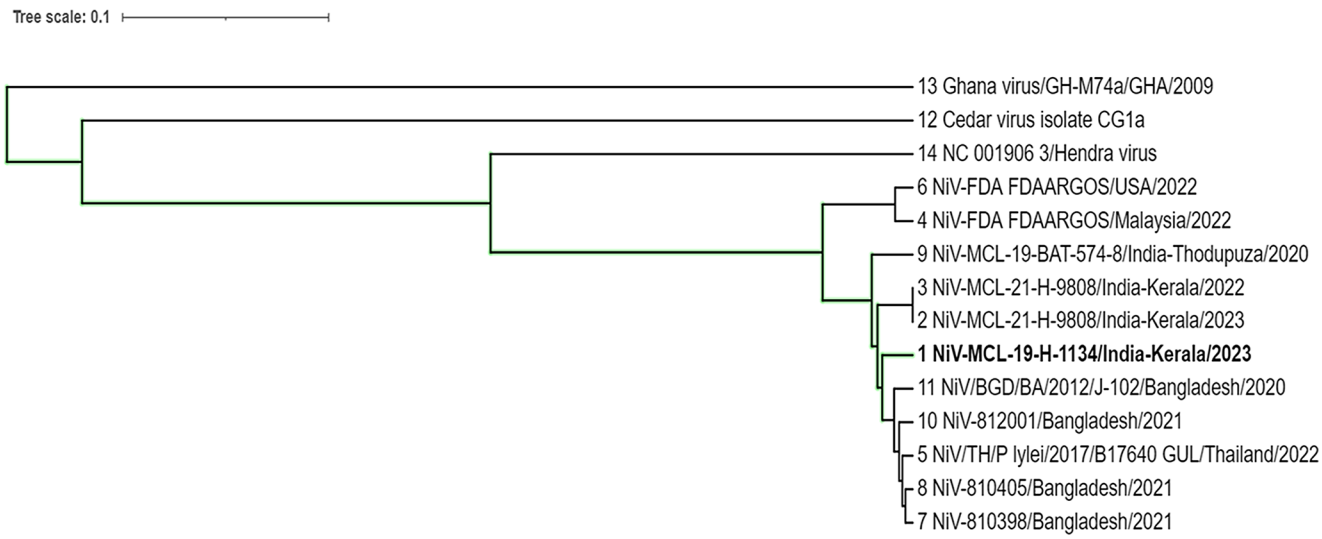

Fig. 2 Phylogenetic tree of selected NiV strains

Molecular docking of MEV and interaction with MHC-I and MHC-II

Molecular docking of formulated vaccines with human MHC-I and MCH-II is done to analyze vaccine candidates' binding affinity and compatibility. MHC-I and MHC-II sequences are retrieved from the PDB server. Both the receptor and ligand PDB files are uploaded to the Clus Pro server (https://cluspro.bu.edu/results.php) to perform molecular docking. They provide ten docked model based on the binding affinity and energy scoring with descending ranking order. The model with the best energy score was downloaded and visualized through PyMol software. The docked PDB file is uploaded in PDBsum to study the amino acid interactions between vaccine and receptor chains. It also provides the type of interactions and the number of residues involved.

Codon optimization analysis and in-silico cloning of vaccine candidate

For the validation of the formulated vaccine, its expression in a host system should be checked. To analyze the optimal expression of the vaccine construct inside the *E.coli* expression system, IDT (Integrated DNA Technology) codon optimization tool (https://www.idtdna.com/pages/tools/codon-optimization-tool?returnurl=%2FCodonOpt) was explored for codon analysis, and optimization. Codon optimization analysis is essential to study the expression of vaccines in a foreign host system. IDT converted our vaccine amino acid sequence into a nucleotide sequence, followed by an optimization according to the *E.coli* (host organism chosen). Codon adaptation index (CAI) and GC content were calculated using the JCAT tool (https://www.jcat.de/). Furthermore, the vaccine cDNA was analyzed for the restriction enzyme; later, the gene of interest (here, the vaccine construct with restriction sites) was inserted in PET28 a (+) plasmid using SnapGene software (www.snapgene.com).

Result

Phylogenomics analysis

The linear RNA genome of NiV MCL-19-H-1134 isolate 2023 strain is visualized in Fig. 2. RNA genome of Paramyxoviridae family along with Henipavirus nipahense MCL-19-H-1134 isolate 2023 strain are aligned to obtain the phylogenetic tree represented diagrammatically in Fig. 3. The Ghana virus, Cedar virus, and Hendra virus were presented as outgroups. MCL-19-H-1134 isolate 2023 strain is closer to NiV-MCL-21-H-9808 India kerala 2022 isolate (no.2) and Bangladesh 2020 isolate (no.11). NiV Malaysia 2022 strain(no.4) and NiV USA 2022 strain (no.6) have common ancestor which lie in a same lineage of Hendra virus (no.14.

Allergen and non-allergen proteins in NiV

Allergens are a category of antigens that elicit a strong immune response, enabling the immune system to combat a particular threat that would otherwise be harmless to the body. These antigens have the potential to trigger type-1 hypersensitivity reactions, leading to allergic responses. Due to their unsuitability for epitope-based vaccine development, we deliberately excluded all allergens from our research and focused exclusively on non-allergenic elements for subsequent investigations. The analysis includes a thorough examination of nine proteins: nucleocapsid protein, phosphoprotein, V protein, W protein, C protein, matrix protein, fusion protein, attachment glycoprotein, and polymerase. With the exception of the C protein, all the other proteins are determined to be non-allergenic, prompting further investigation.

Antigen and non-antigen protein analysis

An antigenic protein, when introduced into a host, can induce a response, making it an ideal candidate for vaccine development. Consequently, all nine proteins were obtained from NCBI and assessed for their antigenicity using the VaxiJen v2.0 server, which assigned an antigenic score to

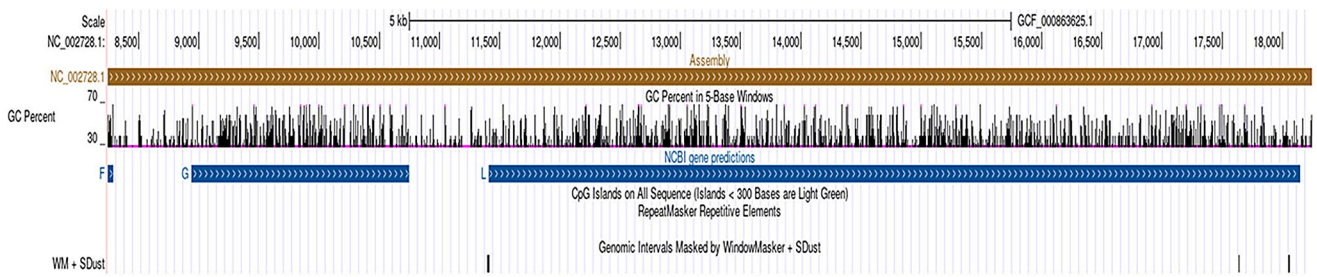

Fig. 3 Linear RNA genome of NiV MCL-19-H-1134 isolates 2023 strain

Table 1 All protein of NiV with their antigen score (threshold 0.7) analysed in VaxiJen v2.0

Proteins	Score	Antigen/ non antigen
Nucleocapsid protein	0.6709	Non-antigen
Phosphoprotein	0.6669	Non antigen
V protein	**0.7219**	**Antigen**
W protein	**0.8053**	**Antigen**
C protein	0.4380	Non antigen
Matrix protein	0.4829	Non-antigen
Fusion protein	0.3644	Non-antigen
Attachment glycoprotein	0.4417	Non-antigen

each protein. A threshold of 0.7 was set; proteins with scores exceeding 0.7 were classified as antigenic, while those below were deemed non-antigenic. Upon scrutinizing the NiV proteins, the VaxiJen scores for V protein (WKR82403.1) and W protein (WKR82404.1) surpassed the 0.7 threshold Table 1. Consequently, further analysis was conducted on V and W proteins, while the remaining proteins were excluded from consideration in vaccine construction.

Cytotoxic T-cell (CTL) epitope prediction

AllergenFp analysis predicts two non-allergenic proteins; V protein (protein id- WKR82403.1) and W protein (protein id- WKR82404.1) as antigenic which fulfil the criterion (antigen score 0.7219 and 0.8053) to run it into the NetCTL 1.2 server for CTL prediction. Among them, W protein has the highest antigen score (Table 1) hence it will create a high immune response in human cells. CTL analysis is done for W protein. Cytotoxic T cells are activated by MCH class I epitopes resulting in the activation of adaptive immune response. For our study epitopes showing a binding score of more than 0.8 are considered as CTL epitope. Where the NetCTL score represents the binding affinity and specificity of the epitope to MHC. Three epitopes VSDAKVLSY, LSDAKVLS, and NSDAQPLYY which are 9-mer sequences having NetCTL scores 0.996544, 0.949801, 0.923207, respectively are found against the HLA-A*01:01 allele. Similarly, SPQKRLPML, MPKSRGIPI and KPADAPGAL are found against HLA-B*07:02 allele, and AIPFTPKNL is found against HLA-C*01:02, Table 2.

Helper T-cell prediction

NiV W protein was analysed to predict the possible epitope that binds more likely to respective MHC-II complexes. The NetMHCIIpan-4.1 server utilizes Artificial Neural Networks (ANNs) to forecast the binding of peptides to MHC II molecules with known protein sequences. According to NetMHCIIpan prediction epitopes with high binding scores are sorted. There are a total of four epitopes (E) showing promising binding affinity with different HLA alleles. E1 (RETDLVHLE) show high binding affinity with allele HLA-DPA10103-DPB10101, HLA-DPA10106-DPB10301, HLA-DQA10101-DQB10201, HLA-DQA10102-DQB10202, HLA-DQA10104-DQB10204. E2 (ITSDAVQNA) bind promisingly with HLA-DQA10101-DQB10201, HLA-DQA10102-DQB10202, HLA-DQA10103-DQB10203, HLA-DQA10104-DQB10204, HLA-DQA10105-DQB10205 alleles. E3(IVGISPEEE) has good binding affinity with HLA-DQA10101-DQB10201, HLA-DQA10102-DQB10202,

Table 2 CTL epitope analysis of W protein of NiV by NetCTL-1.2 server

Allele	Start	End	Length	Peptide	Core	Icore	Score	Percentile rank
HLA-A*01:01	151	159	9	VSDAKVLSY	VSDAKVLSY	VSDAKVLSY	0.996544	0.01
HLA-A*01:01	150	159	10	LVSDAKVLSY	LSDAKVLSY	LVSDAKVLSY	0.949801	0.02
HLA-A*01:01	353	365	13	NSQQGKDAQPLYY	NSDAQPLYY	NSQQGKDAQPLYY	0.923207	0.03
HLA-A*01:01	106	116	11	QLDPVVTDVVY	QLDPTDVVY	QLDPVVTDVVY	0.83942	0.05
HLA-A*01:01	148	159	12	VCLVSDAKVLSY	VSDAKVLSY	VCLVSDAKVLSY	0.780317	0.07
HLA-A*02:01	106	114	9	QLDPVVTDV	QLDPVVTDV	QLDPVVTDV	0.870003	0.05
HLA-A*02:01	155	163	9	KVLSYAPEI	KVLSYAPEI	KVLSYAPEI	0.600046	0.2
HLA-A*02:01	105	114	10	IQLDPVVTDV	IQLDPVVTV	IQLDPVVTDV	0.582264	0.21
HLA-A*02:01	207	215	9	VIAEHYYGL	VIAEHYYGL	VIAEHYYGL	0.553093	0.23
HLA-A*02:01	105	115	11	IQLDPVVTDVV	IQLDPVVTV	IQLDPVVTDVV	0.475095	0.28
HLA-B*07:02	320	328	9	SPQKRLPML	SPQKRLPML	SPQKRLPML	0.972555	0.02
HLA-B*07:02	396	404	9	MPKSRGIPI	MPKSRGIPI	MPKSRGIPI	0.939144	0.04
HLA-B*07:02	302	311	10	KPADAPGAGL	KPADAPGAL	KPADAPGAGL	0.932132	0.04
HLA-B*07:02	394	402	9	TPMPKSRGI	TPMPKSRGI	TPMPKSRGI	0.892711	0.05
HLA-B*07:02	278	287	10	KPIESVGHIL	KPIESGHIL	KPIESVGHIL	0.850321	0.06
HLA-C*01:02	190	198	9	AIPFTPKNL	AIPFTPKNL	AIPFTPKNL	0.904462	0.02
HLA-C*01:02	319	327	9	KSPQKRLPM	KSPQKRLPM	KSPQKRLPM	0.552218	0.07
HLA-C*01:02	193	200	8	FTPKNLSV	FTP-KNLSV	FTPKNLSV	0.516517	0.08
HLA-C*01:02	199	207	9	SVPAKDSPV	SVPAKDSPV	SVPAKDSPV	0.420344	0.11
HLA-C*01:02	372	379	8	RSPDKTEI	RSPDK-TEI	RSPDKTEI	0.349253	0.14

HLA-DQA10103-DQB10203, HLA-DQA10104-DQB10204, HLA-DQA10105-DQB10205 alleles. And E4 (YRSIEGSRS) with DRB1_0101, DRB1_0103, DRB1_0105 alleles. Protein sequence with core epitope and their binding HLA alleles are listed in (online resource 1).

Identification of B-cell epitope

B-lymphocytes play a crucial role in the humoral immune system, producing a diverse array of pathogen-specific antibodies that not only reduce the viral load but also contribute to antigen neutralization. To predict epitopes capable of activating B-cells, the antigenic proteome was assessed using the IEDB server. A threshold of 0.500 was considered (Fig. 4) and total of six epitopes (online resource 2), varying in length, were identified as having the potential to activate B-cells and were consequently included in this study.

Population coverage analysis

While studying population coverage analysis of the epitopes and their corresponding MHC alleles, we received three major epitopes interacting with alleles that have the most population coverage throughout the globe. Population of a total of 14 countries are studied from them RETDLVHLE, KNLSVPAKD, and AIPFTPKNL shown creating immune response in populations. Among the above RETDLVHLE is the highest, it covers 95.63% population in Brazil, 87.61% of China, 79.86% of England, 94.86% of India, 98.04% of Russia, and covers 99.94% population in the United States. Epitopes with their population coverage for all the countries are given in the Fig. 5.

Toxicity, immunogenicity, and physiochemical analysis

Antigenic epitopes are further analysed through ToxinPred server. Epitopes with their toxicity prediction score are given in Table 3, only the non-toxic epitopes were considered for further study. Along with the toxin prediction and immunogenicity other physiochemical properties like Hydrophobicity, Hydropathicity, Hydrophilicity, Charge, and Molecular weight are also considered and listed in Table 3. A high immunogenicity score indicates more immunogenic epitopes.

Molecular interaction and validation of epitope

Molecular docking analysis of CTL and HTL with HLA allele

Molecular docking study was conducted to study the interaction of epitopes with different HLA alleles. The study revealed that a total of four epitopes have high binding affinity to the HLA allele which are RETDLVHLE, KNLSVPAKD, YRSIEGSR, and AIPFTPKNL. Epitope RETDLVHLE binds with high affinity with HLA-DP as well as HLA-DQ with a binding energy of -15.6 kcal/mol and −14.3 kcal/mol respectively. Epitopes with their binding affinity with respective HLA alleles are given in Table 4. Molecular docking interactions of epitopes with HLA alleles and their visualization followed by amino acid interaction are shown in Figs. 6, 7, 8, 9 and 10.

Molecular interaction analysis of epitopes with TLR-3

TLR protein plays an important role in the recognition of pathogens and stimulates various signaling pathways in the viral infection (Christmas 2010). A molecular interaction

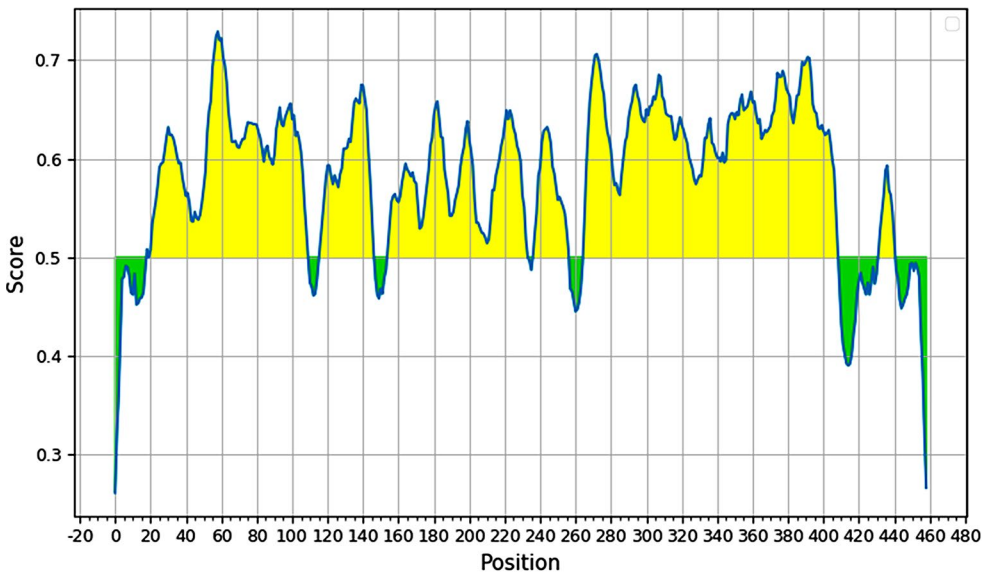

Fig. 4 B-cell epitope prediction by IEDB server. The line above the threshold (yellow) represents B-cell epitope and the line below the threshold (green) does not. Average: 0.582 Minimum: 0.261 Maximum: 0.729 Center position: 4 Threshold: 0.500

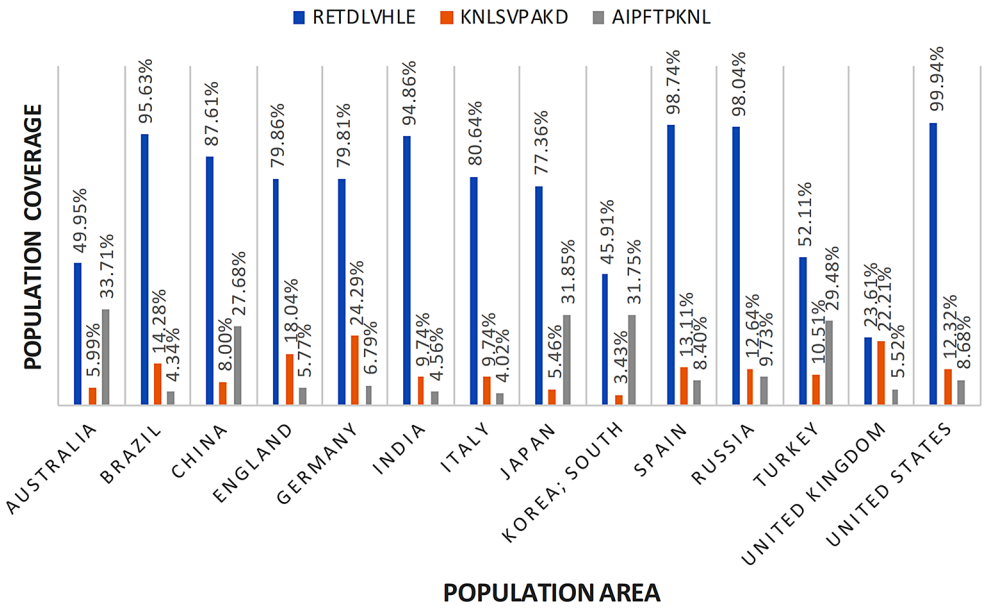

Fig. 5 Population conservancy analysis of epitope RETDLVHLE, KNLSVPAKD, and AIPFTPKNL across different countries of the world

Table 3 Toxicity of epitopes with SVM score and other chemical properties

Peptide sequence	SVM score	Prediction	Hydrophobicity	Hydropathicity	Hydrophilicity	Charge	Mol wt.
AIPFTPKNL	-0.95	Non-Toxin	0.01	0.18	-0.42	1.00	1000.33
IVGISPEEE	-0.98	Non-Toxin	-0.00	-0.01	0.47	-3.00	972.19
KNLSVPAKD	-0.76	Non-Toxin	-0.29	-0.82	0.63	1.00	971.24
KPADAPGAL	-0.86	Non-Toxin	-0.06	-0.20	0.30	0.00	839.07
KSPQKRLPM	-0.78	Non-Toxin	-0.47	-1.57	0.71	3.00	1084.47
LQIKGNKPA	-0.61	Non-Toxin	-0.21	-0.74	0.26	2.00	968.30
MPKSRGIPI	-1.21	Non-Toxin	-0.15	-0.21	0.16	2.00	998.38
QLDPTDVVY	-1.30	Non-Toxin	-0.07	-0.19	-0.13	-2.00	1212.46
RETDLVHLE	-0.46	Non-Toxin	-0.30	-0.79	0.67	-1.50	1111.35
RETDLVHLE	-0.46	Non-Toxin	-0.30	-0.79	0.67	-1.50	1111.35
SPQKRLPML	-0.99	Non-Toxin	-0.29	-0.71	0.18	2.00	1069.46
VIAEHYYGL	-0.20	Non-Toxin	0.12	0.46	-0.77	-0.50	1177.52
VSDAKVLSY	-1.05	Non-Toxin	-0.05	0.37	-0.10	0.00	1144.42
YRSIEGSR	-0.57	Non-Toxin	-0.47	-1.41	0.69	1.00	967.15
YTSDDEEAD	-0.87	Non-Toxin	-0.40	-2.06	1.34	-5.00	1044.05

Table 4 Binding affinity of different epitopes with their subsequent HLA alleles. Epitope with higher binding affinity is highlighted

MHC	Allele	Sequence	Epitope	Binding energy (kcal/mol)
MHC II	**HLA-DPA10103-DPB10101**	**SKEDRETDLVHLEDK**	**RETDLVHLE**	**-15.6**
MHCII	**HLA-DPA10106-DPB10301**	**FTPKNLSVPAKDSPV**	**KNLSVPAKD**	**-13.4**
MHC II	HLA-DPA10104-DPB10201	DSPVIAEHYYGLGVR	VIAEHYYGL	-8.9
MHC II	HLA-DQA10101-DQB10201	DSIKLYTSDDEEADQ	YTSDDEEAD	-12
MHC II	**HLA-DQA10101-DQB10201**	**SKEDRETDLVHLEDK**	**RETDLVHLE**	**-14.3**
MHC II	HLA-DQA10101-DQB10201	SEVIVGISPEEEEPS	IVGISPEEE	-11.6
MHC II	**DRB1_0101**	**PLYYRSIEGSRSPDK**	**YRSIEGSR**	**-12.1**
MHC II	DRB1_0103	RDSLQIKGNKPADAP	LQIKGNKPA	-9.2
MHCI	HLA-B*07:02	SPQKRLPML	SPQKRLPML	-11.3
MHCI	HLA-B*07:02	MPKSRGIPI	MPKSRGIPI	-10.5
MHCI	HLA-B*07:02	KPADAPGAGL	KPADAPGAL	-7.5
MHCI	HLA-C*01:02		KSPQKRLPM	-11.6
MHCI	**HLA-C*01:02**		**AIPFTPKNL**	**-13.1**
MHCI	HLA-A*01:01		VSDAKVLSY	-11.5
MHCI	HLA-A*01:01		QLDPTDVVY	-11.2

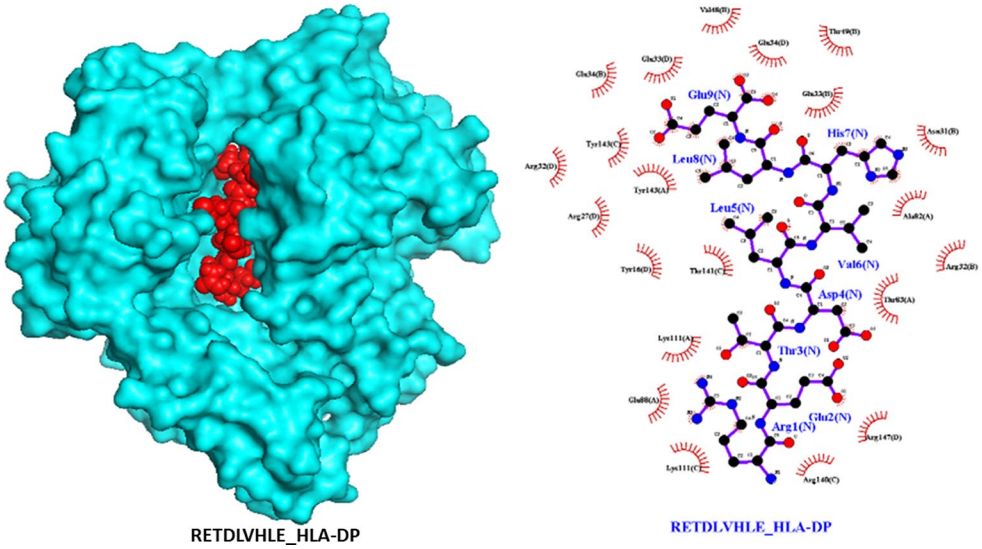

Fig. 6 Interaction of RETDLVHLE with HLA DP and amino acid interaction is visualized through Ligplot software

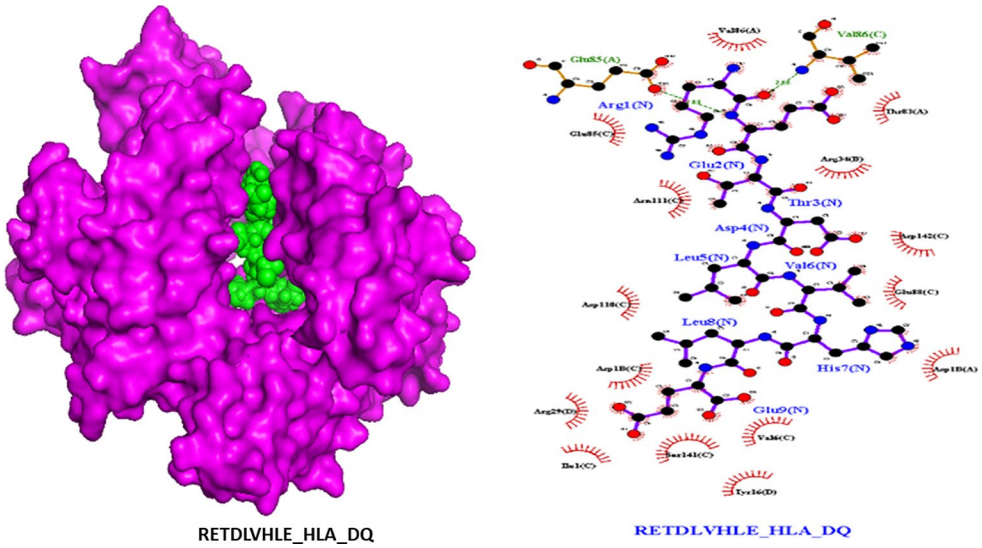

Fig. 7 Interaction of RETDLVHLE with HLA DQ and amino acid interaction is visualized through Ligplot software

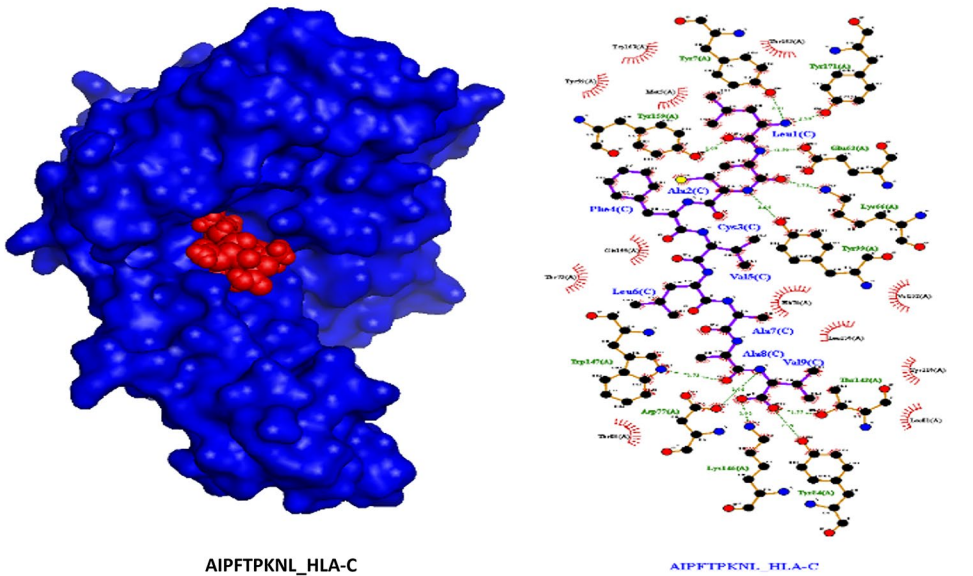

Fig. 8 Interaction of AIPFTPKNL with HLA C and amino acid interaction is visualized through Ligplot software

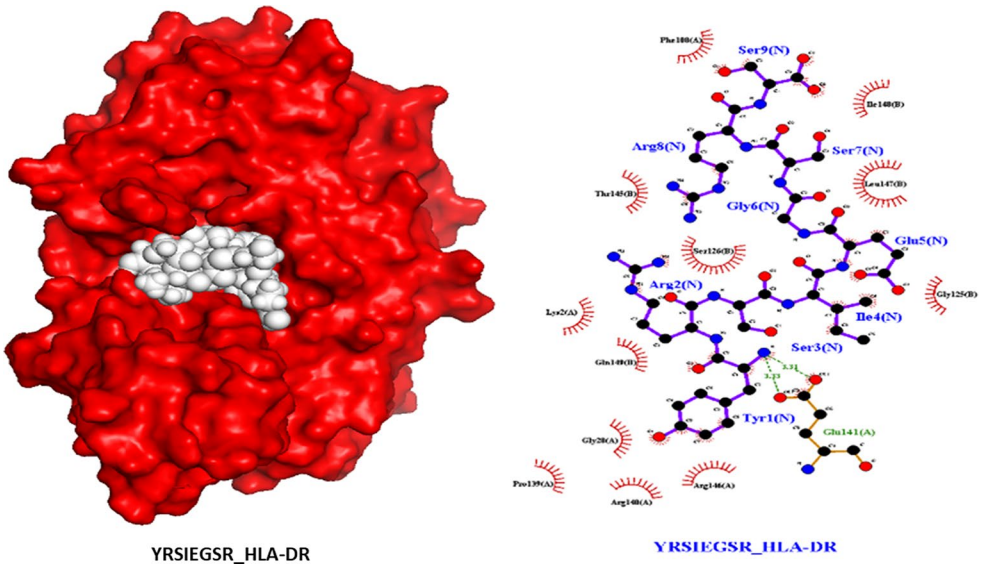

Fig. 9 Interaction of YRSIEGSR with HLA DR and amino acid interaction is visualized through Ligplot software

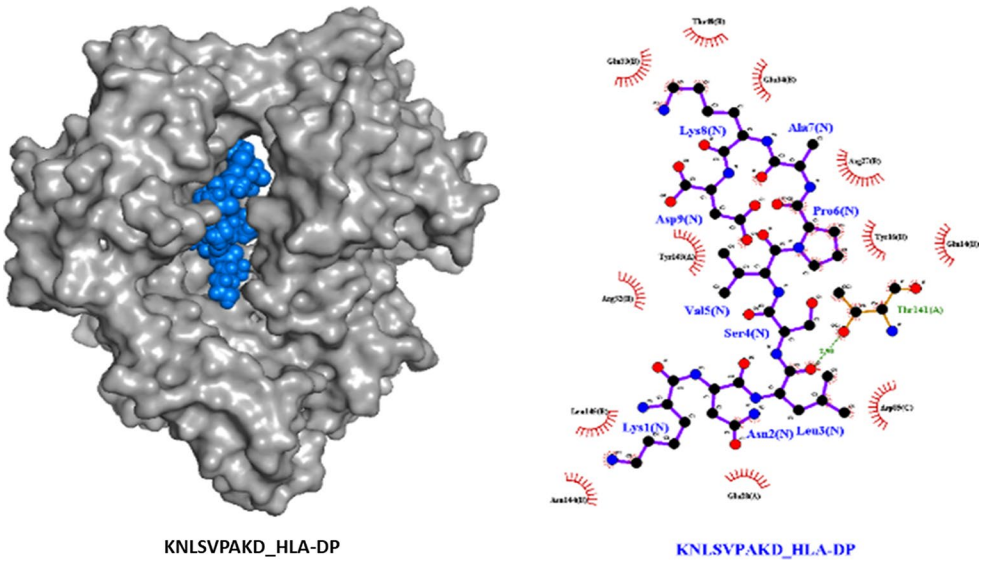

Fig. 10 Interaction of KNLSVPAKD with HLA DP and amino acid interaction is visualized through Ligplot software

study of TLR3 with the epitopes of the NiV was done. A binding affinity of -11.6 kcal/mol, -10.5 kcal/mol, -10.1 kcal/mol, and −9.5 kcal/mol with epitope RETDLVHLE, KNLSVPAKD, AIPFTPKNL and YRSIEGSR respectively was recorded. The interaction of all the epitopes with TLR3 protein is shown in (online resource 3).

Multi-epitope vaccine designing and characterization

Multi-epitope vaccine formulation

Six selected epitopes from CTLs, HTLs, and B cell epitopes are joined with appropriate linker sequences (AAY, GGGGS, and KK) and aligned in a straight line. EAAAK linkers are used to join two consecutive epitopes with specific adjuvants. A schematic representation of the MEV construct is shown in Fig. 12b.

Properties of the constructed vaccine candidate

The final vaccine construct consists of 480 amino acids, verified by the result obtained from the ProtPharam tool (https://web.expasy.org/protparam/) analysis. Vaxijen 2.0 and AllergenFP web server results indicate that the vaccine construct is antigenic and non-allergenic. Furthermore, the molecular weight of the vaccine is 50721.39, and the molecular formula is $C_{2248}H_{3536}N_{604}O_{699}S_{16}$. Other physiochemical properties like extinction coefficient, theoretical pI, Aliphatic index (AI), instability index (II), and GRAVY are listed in the online resource 4.

In-silico immune simulation of MEV

C-ImmSim result shows the effective mammalian immune response against the formulated vaccine. Figure 11 depicts the computational analysis of the immune response triggered by our vaccine. It is a 3-dose vaccine given at a time interval of 30 days and observed up to 100 days. Upon analyzing the effects of the first dose in contrast to the subsequent second and third doses, it becomes clear that there is a rise in the levels of different antibodies, such as IgM + IgG, IgM, IgG1 + IgG2, IgG1, and IgG2 (Fig. 11a). This suggests that administration of the experimental vaccine results in an enhanced antibody reaction.

Moreover, the Nipah MEV candidate can stimulate the production of different cytokines, such as IFN-gamma, interleukin-4 (IL4), interleukin-12 (IL12), and transforming growth factor-beta (TGF-β). Following the initial dose, the second dose of the vaccine leads to reduced levels of IFN-gamma and TGF beta, while IL-12 remains consistent with no significant alterations (Fig:11b). A subsequent increase was seen in B, TH, and PLB cell populations (Fig. 11c, d, h). Moreover, the population of TC cells starts decreasing 15 days after the first dose, but TC memory cells remain constant (Fig. 11e). Natural killer (NK) cell proliferation was also noted, along with dendritic cells, which play a crucial role as mediators in activating T cells (Peterson and Barry 2021) (Fig. 11f). Dendritic cells maintain an active state throughout the entire duration and do not exhibit significant fluctuations during administering injection doses (Fig. 11g). The number of B cell populations in the active state increases with each vaccine dose (Fig. 11i); likewise, the number of TC cell populations in the active state is maximum in 30 days after the first dose, and on a contradictory resting cell number decreases drastically after the initial vaccination dose (Fig. 11j). Active TH cell and

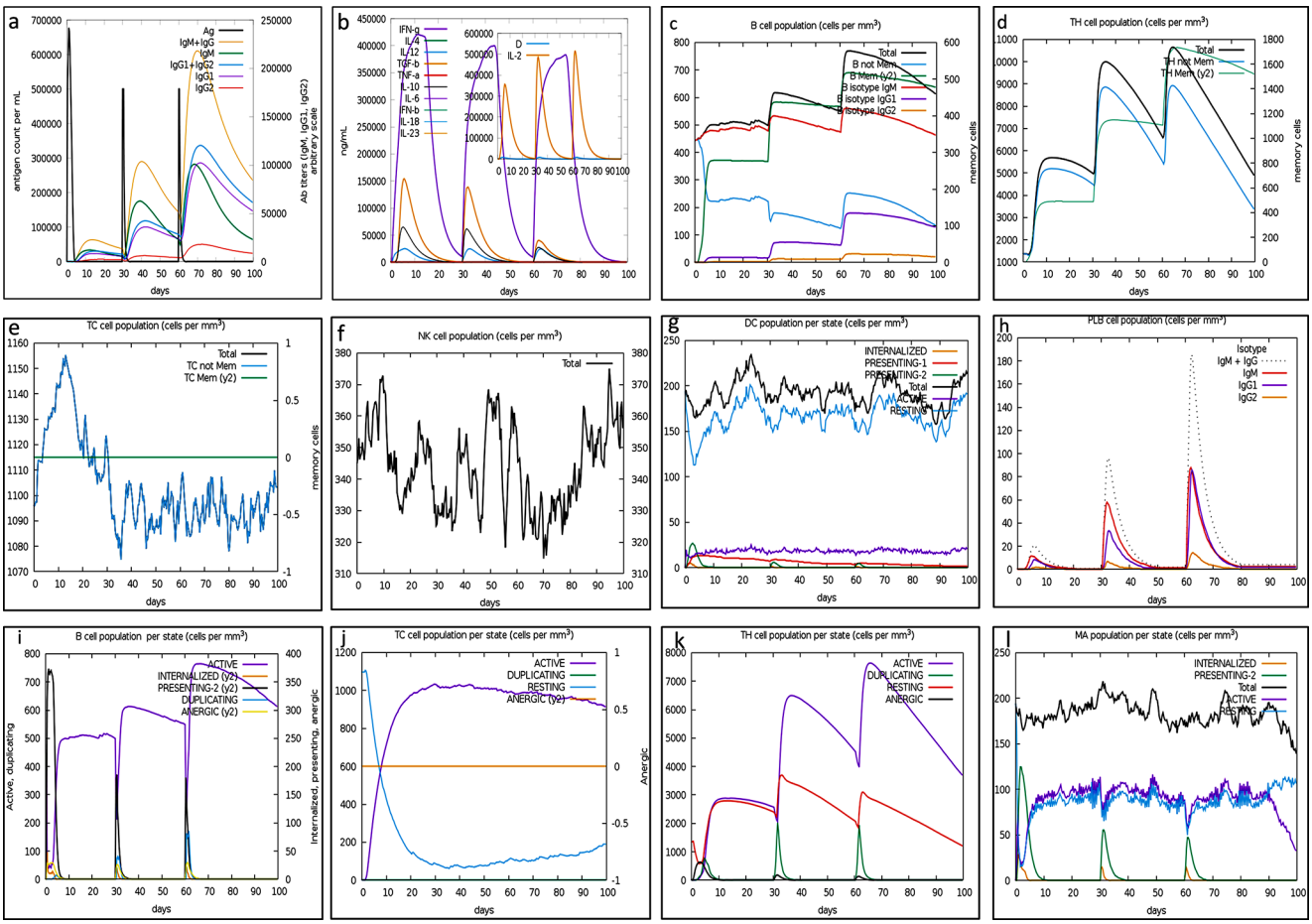

Fig. 11 NiV vaccine immune simulation. (**a**) Antigen and immunoglobulins (**b**) Concentration of cytokines and interleukins. D in the inset plot is danger signal. (**c**) Plasma B lymphocytes count subdivided per isotype (**d**) CD4 T-helper lymphocytes count. The plot shows total and memory counts. (**e**) CD8 T-cytotoxic lymphocytes count. Total and memory shown. (**f**) Natural Killer cells (total count). (**g**) Dendritic cells; The curves show the total number broken down to active, resting, internalized and presenting the ag. (**h**) Plasma B lymphocytes count sub-divided per isotype (**i**) B lymphocytes population per entity-state (**j**) CD8 T-cytotoxic lymphocytes count per entity-state (**k**) CD4 T-helper lymphocytes count sub-divided per entity-state (**l**) MA population per state

MA cell population increases with each vaccine dose shown in Fig. 11k, l.

Structure prediction, refinement, and stability of MEV candidate

The tertiary structure of the MEV construct is predicted and visualized (Fig. 12a). For the stability assessment, the PDBsum server gives the secondary structure of the vaccine with helices stranded by sheets. Figure 12c represents the vaccine secondary structure with beta turns, gamma turns, and beta hairpins. The topology of the vaccine structure is given in Fig. 12d. The Ramachandran plot depicts the most favoured regions containing 92.7% of amino acids Fig. 12e.

Molecular docking of MEV and interaction with MHC-I and MHC-II

Molecular docking of MEV with MHC-I and MHC-II was done in the ClusPro web server. The most suitably docked model was collected and visualized in Fig. 13a and b. The amino acid interactions between ligand and receptor were also analyzed in the PDBsum server. Side chain interactions, represented by coloured lines in the Fig. 14, demonstrated various types of interactions, including salt-bridge (red line), hydrogen bonds (blue lines), and non-bonded contacts (orange dashed lines). In MHC-I, only one side chain (chain D) was found interacting with vaccine residue (chain C) involving 30 amino acids (Fig. 14a), whereas MHC-II interacts by two side chains (chain D and B) with vaccine residue (chain C) involving 28:31 and 1:1 amino acid residue (Fig. 14b, c). Ramachandran plot analysis of docked candidates depicts the most favoured regions containing 81.6% of amino acids for the MEV-MHCI complex (Fig. 14d) and 84.3% of amino acids for the MEV-MHCII complex Fig. 14e.

Codon optimization analysis and in-silico cloning of vaccine candidate

To effectively produce and purify the multi-subunit vaccine, it is essential to have a suitable expression system that

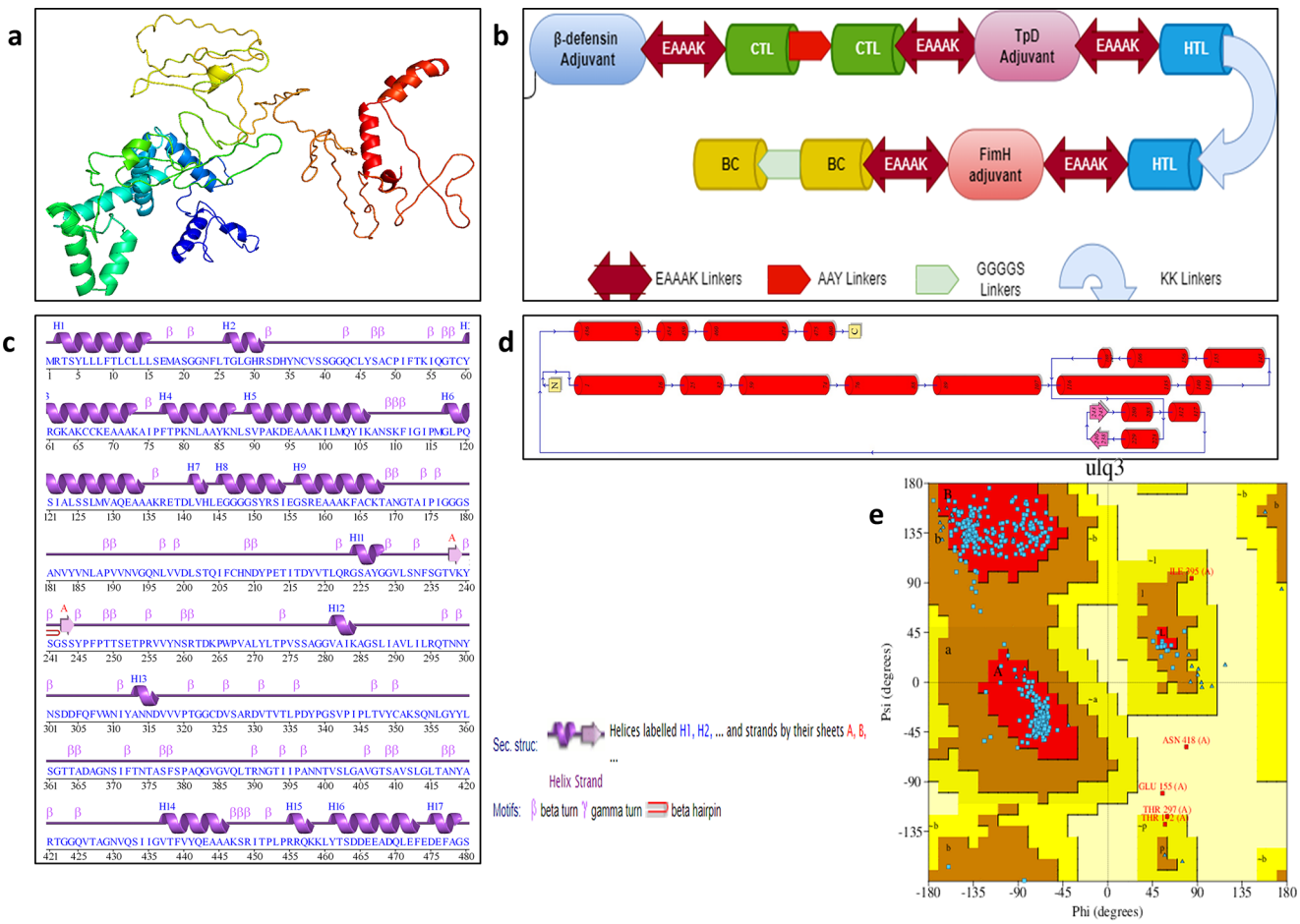

Fig. 12 Structure prediction and validation of MEV candidate. (**a**) Tertiary structure od MEV. (**b**) Schematic representation of MEV formulation. (**c**, **d**) Secondary structure and topology of MEV candidate predicted by PDBsum server. (**e**) Ramachandran plot of MEV tertiary structure

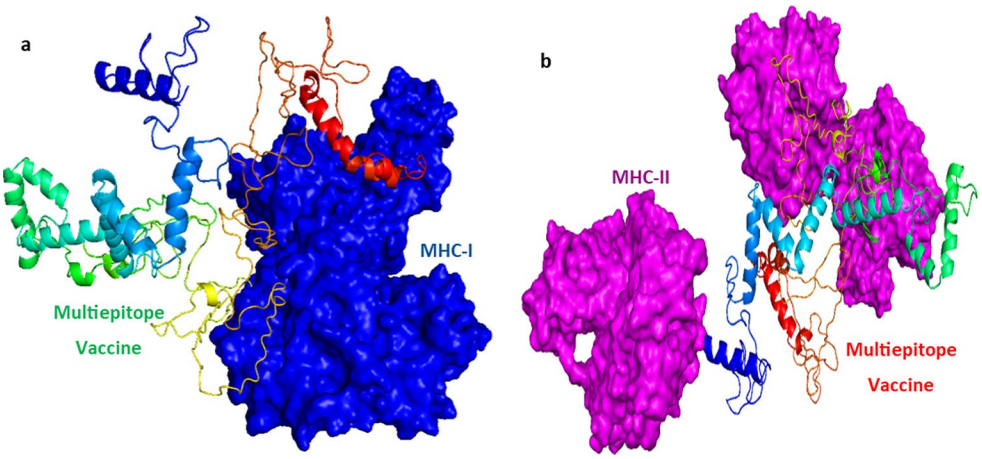

Fig. 13 Molecular docking of MEV. (**a**) MEV-MHC class I complex. (**b**) MEV-MHC class II complex

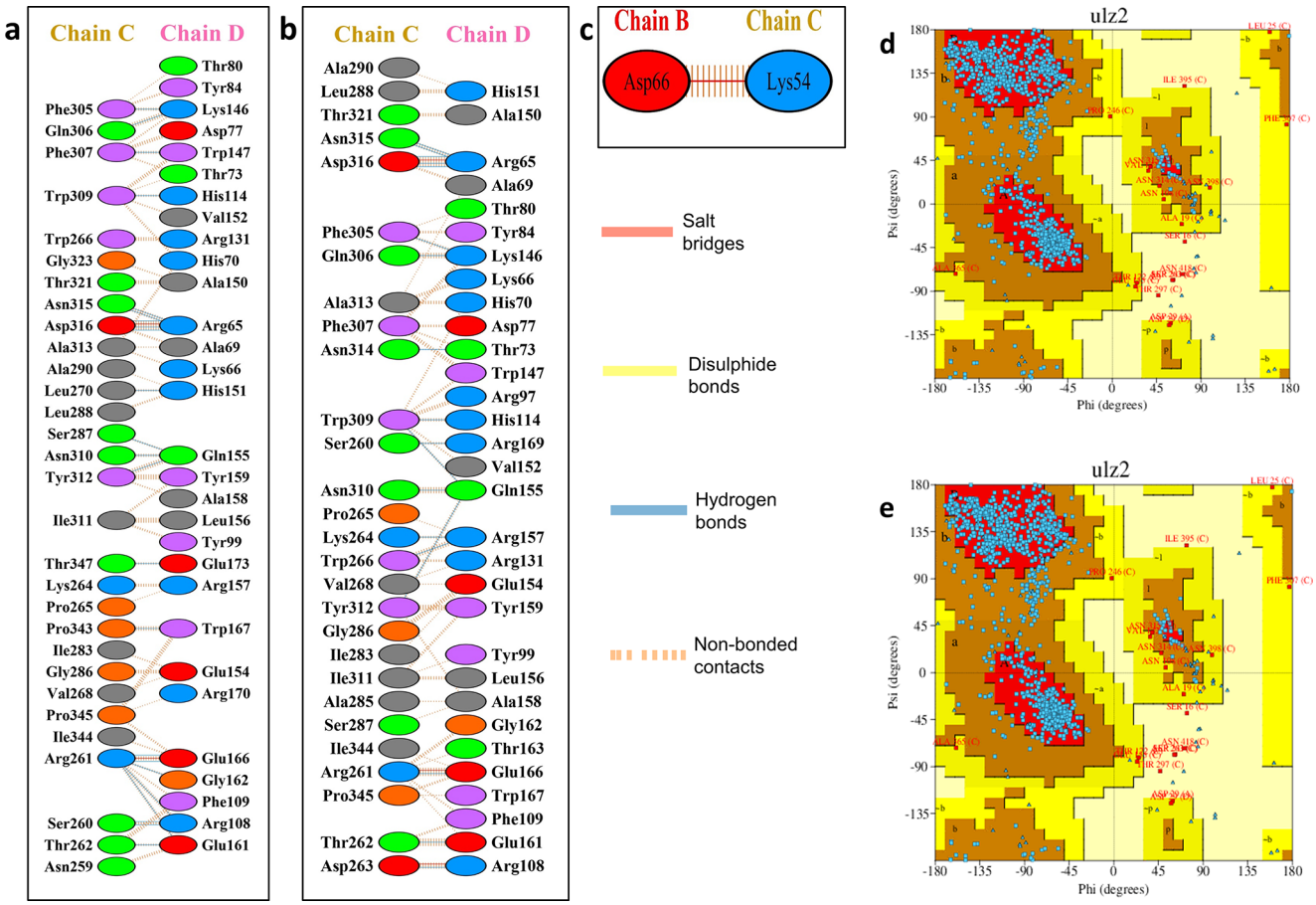

Fig. 14 (**a**) Amino acid residue interaction between MHC-I (chain D) and vaccine (chain C). (**b**, **c**) Amino acid interaction between MHC-II (chain D, B) and vaccine (chain C). (**d**) Ramachandran plot of MEV-MHCI complex tertiary structure. (**e**) Ramachandran plot of MEV-MHC-II complex tertiary structure

incorporates codons specifically chosen to match the expression host (Majee et al. 2021). From 180 amino acid vaccine candidates, 1440 nucleotide cDNA chains were made and optimized according to the *E. coli* host. The selection of a particular codon for coding an amino acid differs depending on the type of host organism (Bahir et al. 2009). Here, typically, the codon adaptation index (CAI) was considered. The CAI value ranges from 0 to 1, creating an array for coding a specific protein by a codon. The higher CAI value of a codon indicates a high bias toward that codon by the organism in translation (Carbone et al. 2003). To enhance codon optimization specifically for the *E.coli* K12 strain, the Java

Codon Adaptation Tool (JCAT) was utilized. In this study, the CAI value of the vaccine construct cDNA for the *E.coli* K12 strain was 0.965, and the GC content was 52.2%. Complete vaccine sequences with linkers and their optimized nucleotide sequence according to the host Organism *E. coli* are provided in the online resource 5. Vaccine cDNA was successfully inserted into the pET28a(+) expression vector with the help of the SnapGene tool; prior to insertion, two restriction sites were inserted respectively (Eco53KI and HpaI) in the start and end region of the vaccine sequence. The in silico cloning creates a 5367 bp expression vector Fig. 15 which can be inserted into the host organism for expression.

Discussion

In the field of in silico vaccine design, commonly referred to as reverse vaccinology, the tasks involve screening, selecting, and mapping vaccines using bioinformatics tools and software (Sharma et al. 2021). In this context, our research evaluates a pan-specific MEV designed for the NiV. NiV, causing acute respiratory infection and fatal encephalitis in humans, had a significant outbreak in Malaysia in 1998. Ribavirin treatment during this outbreak led to a 36% reduction in mortality for patients with NiV encephalitis (Chong et al. 2001; Johnson et al. 2021). Although the practice of self-medication could exacerbate the vulnerability of the situation (Zeb et al. 2022). Recently, September 2023 recorded several deaths caused by the NiV in Kerala, India, and brought the virus back into focus. Despite this, no commercial vaccines are available for the virus Hence, significant approaches should be obligatory to build active immunity and approved therapeutics for the NiV. For the management of NiV outbreak, approved therapeutics are needed, and, in that case, epitope-subunit vaccines derived from pathogen proteome are the best approach. In the past, numerous investigations have been done on the Nipah structural proteins, particularly Glycoprotein G and phosphoprotein proteins, as targets for creating vaccine candidates through various approaches. (Kaushik 2020; Majee et al. 2021; Srivastava et al. 2023).

In this study, phylogenetic analysis was done, which revealed the relative closeness between viruses of the same genus; representative strains were selected for the development of a vaccine. Proteome analysis of 9 NiV proteins revealed that V protein (WKR82403.1) and W protein (WKR82404.1) are non-allergic and antigenic; among them, W protein has the maximum antigen score (Table 1). So, the server prediction for the W protein makes it more acceptable, and the epitopes from the W protein will cause a high immune response. As indicated in Table 4, the peptides YRSIEGSR and AIPFTPKNL exhibit substantially negative docking scores (-12.1 and −13.1 kcal/mol, respectively) when interacting with MHC class I alleles, suggesting high binding affinities. Similarly, the peptides RETDLVHLE and KNLSVPAKD show more negative docking scores with MHC class II complexes, indicating strong binding affinities to these MHC alleles. Comparable results were observed for two B cell epitopes. Also, they show high population coverage around the globe, which can create an immune response in a wide range of populations. The epitope RETDLVHLE

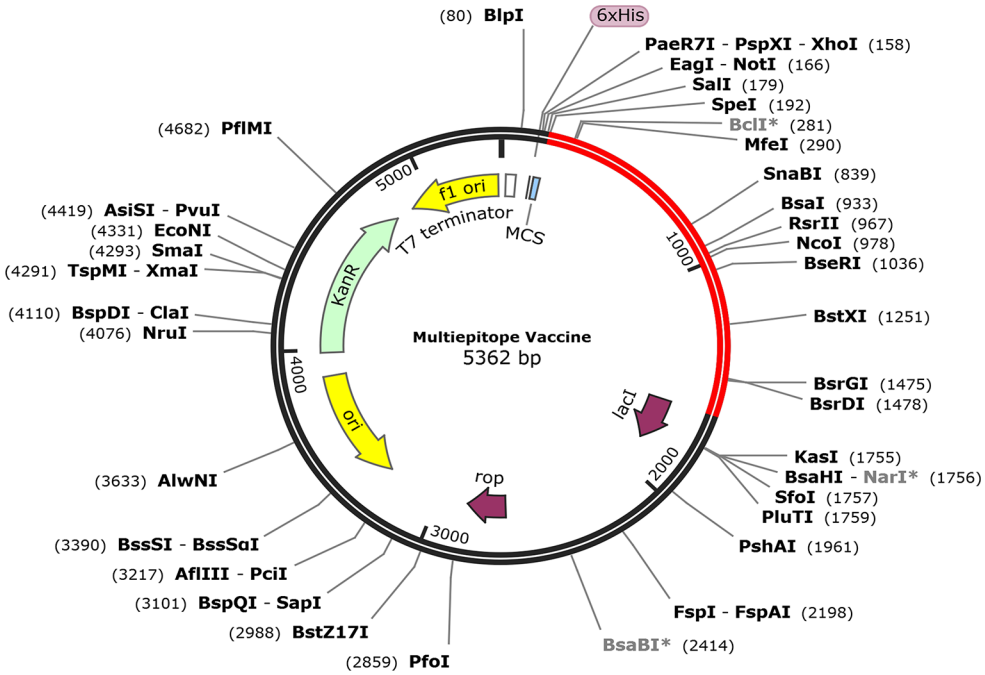

Fig. 15 In silico cloning of formulated vaccine in pET28a(+)

binds to TLR-3 with an affinity of -11.6 kcal/mol and also demonstrates strong interaction with the TAP cavity, highlighting its potential effectiveness for vaccine development.

The vaccine consists of 480 amino acids along with necessary linker sequences. The vaccine construct is linked with β-defensin, universal TpD, and the final subunit of *E.coli* type 1 fimbria (FimH) adjuvants to boost the immune response. Prior investigations have shown that the *E. coli* type 1 FimH adjuvant induces the production of TNF-α and IFN-γ, promoting the proliferation of local dendritic cells (Zhang et al. 2022). Conversely, TpD has demonstrated the capability to stimulate the production of neutralizing antibodies (Chan et al. 2020; Li et al. 2018). With a negative GRAVY score of -0.009, indicating hydrophilicity, and an instability index of 32.49, below the threshold of 40, denoting stability, our vaccine candidate exhibits favorable physicochemical characteristics. Similar findings were observed in a study by Nguyen, T. L. and Kim, H. (2024) against the Powassan virus. In in silico immune simulation study, successive vaccine doses led to a notable increase in the concentration of IgM + IgG, IgM, IgG1 + IgG2, IgG1, and IgG2, indicating a robust immune response (Fig. 11a). Molecular docking was employed to study the affinity between the vaccine and human major histocompatibility complexes. The 3D structure of the molecular docked complexes was visualized in Fig. 12a, and Ramachandran plot results confirmed the stability of the docked complex. In the MEV-MHCI docked complex, 81.6% of amino acids resided in the most favoured region (Fig. 14a), while in the MEV-MHCII docked complex, 84.3% of amino acids were within the most favored region (Fig. 14b). Side-chain interactions, represented by coloured lines in Fig. 14a, b,c, signify the effective molecular complex formation between MEV and MHC. Overall, our MEV candidate exhibits high stability in terms of structural conformation, favorable physiochemical characteristics, strong molecular bonding with MHC, and, importantly, a robust immune response in the human body.

To check the adaptation of the vaccine, the amino acid sequence is optimized as per the expression host system (*E. coli* K12). A 1440 bp host-specific optimized cDNA with 52.2% GC content and a CAI value of 0.965 indicates a high level of adaptation of a gene's codon usage to the preferred codon usage of *E. coli* is inserted into pET28a(+) expression vector for in-silico cloning. Similar work has been done by Nayak et al. 2024 on monkeypox virus disease.

However, designing a MEV using bioinformatics has significant limitations. As dynamic macromolecular complexes, proteins can exhibit unexpected interactions within biological systems. Moreover, reliance on 9-mer epitope prediction methods may overlook significant responses, and combining 9-mer and 10-mer predictions still poses limitations due to variable epitope lengths. Additionally, the promiscuity of HLA class I binding and the poorly defined motifs for less-studied HLA class I alleles exacerbate these challenges (Silva-Arrieta et al. 2020).

Developing an accurate model for such epitopes could advance epitope-based vaccine design. Machine learning-based models show promise for enhancing mRNA vaccine design. Incorporating epitope-based design into mRNA vaccine development could significantly boost their immunogenicity and efficacy. This may be followed by validation in in vitro and in vivo models.

Conclusion

Recently, the viral pandemic in the world has caused serious damage to society, overwhelming healthcare systems and creating unprecedented challenges in managing both critical and routine medical care. The preset study is also related to NiV. The study employs a fundamental immunoinformatic approach to develop a novel MEV for the NiV. By thoroughly analysing NiV structural proteins, we identified key CTLs, HTLs, and B cell epitopes. These epitopes were further screened based on their binding affinity to MHC-HLA alleles. Selected epitopes were utilized to construct a vaccine candidate with significant antigenic potency while being non-toxic, non-allergenic, and possessing favourable physicochemical properties. Molecular docking and immune simulations provided evidence for the efficacy and safety of the formulated vaccine candidate. This computational research substantially contributes to the field and paves the way for in vitro and in vivo studies, ultimately leading to the development of a NiV vaccine. Additionally, our results indicate that the multi-peptide vaccine can be efficiently expressed in a common bacterial system, such as *E.coli*, making it a highly effective prophylactic solution for Nipah viral disease. This vaccine could be tested and potentially deployed in real-world scenarios, supporting global efforts to combat and mitigate the emerging threat of NiV infections.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s40203-024-00246-9.

Acknowledgements Tapas Das, Sutapa Datta, and Arnab Sen acknowledge the University of North Bengal. The authors also thank the Department of Botany and Bioinformatics Facility Centre University of North Bengal.

Author contributions T.D. Conceptualization, Investigation, Methodology, Analysis, data maintenance, data interpretation, Wrote the entire manuscript. S.D. Conceptualization, methodology, Analysis, editing, and data interpretation. A.S. Conceptualization, Methodology, reviewing and editing, formal analysis, and supervision.

Funding This study did not obtain dedicated funding from public, commercial, or non-profit entities.

Data availability No datasets were generated or analysed during the current study.

Declarations

Competing interests The authors declare no competing interests.

References

Abele R, Tampé R (2004) The ABCs of immunology: structure and function of TAP, the transporter associated with antigen processing. Physiology 19(4):216–224. https://doi.org/10.1152/physiol.00002.2004

Alexander J, Sidney J, Southwood S, Ruppert J, Oseroff C, Maewal A, Snoke K, Serra HM, Kubo RT, Sette A, Grey HM (1994) Development of high potency universal DR-restricted helper epitopes by modification of high affinity DR-blocking peptides. Immunity 1(9):751–761. https://doi.org/10.1016/S1074-7613(94)80017-0

Al Naqbi H, Mawart A, Alshamsi J, Al Safar H, Tay GK (2021) Major histocompatibility complex (MHC) associations with diseases in ethnic groups of the Arabian Peninsula. Immunogenetics 73:131–152. https://doi.org/10.1007/s00251-021-01204-x

Aqsha ZM, Dharmawan MA, Kharisma VD, Ansori ANM, Sumantri NI (2023) Reverse vaccinology analysis of B-cell Epitope against Nipah Virus using Fusion protein. Jordan J Pharm Sci 16:499–507. https://doi.org/10.35516/jjps.v16i3.1602

Bahir I, Fromer M, Prat Y, Linial M (2009) Viral adaptation to host: a proteome-based analysis of codon usage and amino acid preferences. Mol Syst Biol 5(1):311. https://doi.org/10.1038/msb.2009.71

Banerjee S, Gupta N, Kodan P, Mittal A, Ray Y, Nischal N, Soneja M, Biswas A, Wig N (2019) Nipah virus disease: a rare and intractable disease. Intractable Rare Dis Res 8(1):1–8. https://doi.org/10.5582/irdr.2018.01130

Bardhan M, Ray I, Roy S, Bhatt P, Patel S, Asri S, Shariff S, Shree A, Mitra S, Roy P, Anand A (2023) Emerging zoonotic diseases and COVID-19 pandemic: global perspective and Indian scenario. Ann Med Surg 85(8):3997–4004. https://doi.org/10.1097/MS9.0000000000001057

Bui HH, Sidney J, Dinh K, Southwood S, Newman MJ, Sette A (2006) Predicting population coverage of T-cell epitope-based diagnostics and vaccines. BMC Bioinformatics 7:1–5. https://doi.org/10.1186/1471-2105-7-153

Calis JJ, Maybeno M, Greenbaum JA, Weiskopf D, De Silva AD, Sette A, Keşmir C, Peters B (2013) Properties of MHC class I presented peptides that enhance immunogenicity. PLoS Comput Biol 9(10):e1003266. https://doi.org/10.1371/journal.pcbi.1003266

Carbone A, Zinovyev A, Képes F (2003) Codon adaptation index as a measure of dominating codon bias. Bioinformatics 19(16):2005–2015. https://doi.org/10.1093/bioinformatics/btg272

Chan Y, Jazayeri SD, Ramanathan B, Poh CL (2020) Enhancement of tetravalent immune responses to highly conserved epitopes of a dengue peptide vaccine conjugated to polystyrene nanoparticles. Vaccines 8(3):417. https://doi.org/10.3390/vaccines8030417

Chapa LG (2023) An Update on Nipah Virus: Past Lessons, Current Efforts, & Future Goals (Doctoral dissertation, University of Pittsburgh)

Chawla M, Cuspoca AF, Akthar N, Magdaleno JSL, Rattanabunyong S, Suwattanasophon C, Jongkon N, Choowongkomon K, Shaikh AR, Malik T, Cavallo L (2023) Immunoinformatics-aided rational design of a multi-epitope vaccine targeting feline infectious peritonitis virus. Front Vet Sci 10. https://doi.org/10.3389/fvets.2023.1280273

Chong HT, Kamarulzaman A, Tan CT, Goh KJ, Thayaparan T, Kunjapan SR, Chew NK, Chua KB, Lam SK (2001) Treatment of acute Nipah encephalitis with Ribavirin. Ann Neurol 49(6):810–813. https://doi.org/10.1002/ana.1062

Christmas P (2010) Toll-like receptors: sensors that detect infection. Nat Educ 3(9):85

Chua KB, Bellini WJ, Rota PA, Harcourt BH, Tamin A, Lam SK, Mahy B (2000) WJ Nipah virus: a recently emergent deadly paramyxovirus. Science 288(5470):1432–1435. https://doi.org/10.1126/science.288.5470.1432

Dallakyan S, Olson AJ (2015) Small-molecule library screening by docking with PyRx. Methods Mol Biol 243–250. https://doi.org/10.1007/978-1-4939-2269-7_19

De Groot AS, Moise L, McMurry JA, Martin W (2009) Epitope-based immunome-derived vaccines: a strategy for improved design and safety. Clin Appl Immunomics 39–69. https://doi.org/10.1007/978-0-387-79208-8_3

Doytchinova IA, Flower DR (2007) VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. BMC Bioinformatics 8:1–7. https://doi.org/10.1186/1471-2105-8-4

Gasteiger E, Hoogland C, Gattiker A, Duvaud SE, Wilkins MR, Appel RD, Bairoch A (2005) Protein identification and analysis tools on the ExPASy server. Humana Press 571–607. https://doi.org/10.1385/1-59259-890-0:571

Gebreyes WA, Dupouy-Camet J, Newport MJ, Oliveira CJ, Schlesinger LS, Saif YM, Kariuki S, Saif LJ, Saville W, Wittum T, Hoet A (2014) The global one health paradigm: challenges and opportunities for tackling infectious diseases at the human, animal, and environment interface in low-resource settings. PLoS Negl Trop Dis 8(11):3257. https://doi.org/10.1371/journal.pntd.0003257

Godden JW, Xue L, Bajorath J (2000) Combinatorial preferences affect molecular similarity/diversity calculations using binary fingerprints and Tanimoto coefficients. J Chem Inf Comput Sci 40(1):163–166. https://doi.org/10.1021/ci990316u

Heo L, Park H, Seok C (2013) GalaxyRefine: protein structure refinement driven by side-chain repacking. Nucleic Acids Res 41:384–388. https://doi.org/10.1093/nar/gkt458

Hsu VP, Hossain MJ, Parashar UD, Ali MM, Ksiazek TG, Kuzmin I, Niezgoda M, Rupprecht C, Bresee J, Breiman RF (2004) Nipah virus encephalitis reemergence, Bangladesh. Emerg Infect Dis 10(12):2082. https://doi.org/10.3201/eid1012.040701

Janeway C, Travers P, Walport M, Shlomchik M (2001) Immunobiology: the immune system in health and disease, vol 2. Garland Pub, New York, p 154

Jespersen MC, Peters B, Nielsen M, Marcatili P (2017) BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. Nucleic Acids Res 45:24–29. https://doi.org/10.1093/nar/gkx346

Johnson K, Vu M, N Freiberg A (2021) Recent advances in combating Nipah virus. Fac Reviews 10. https://doi.org/10.12703/r/10-74

Kaushik V (2020) In silico identification of epitope-based peptide vaccine for Nipah virus. Int J Pept Res Ther 26:1147–1153. https://doi.org/10.1007/s10989-019-09917-0

Khandia R, Singhal S, Kumar U, Ansari A, Tiwari R, Dhama K, Das J, Munjal A, Singh RK (2019) Analysis of Nipah virus codon usage and adaptation to hosts. Front Microbiol 10:439603. https://doi.org/10.3389/fmicb.2019.00886

Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A (2019) RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. Bioinformatics 35:4453–4455. https://doi.org/10.1093/bioinformatics/btz305

Larsen MV, Lundegaard C, Lamberth K, Buus S, Lund O, Nielsen M (2007) Large-scale validation of methods for cytotoxic

T-lymphocyte epitope prediction. BMC Bioinformatics 8:1–12. https://doi.org/10.1186/1471-2105-8-424

Letunic I, Bork P (2021) Interactive tree of life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res 49:293–296. https://doi.org/10.1093/nar/gkab301

Li X, Zhang S, Lei J, Zhu Y, Zhou X, Xiao J, Xiang T (2018) Prophylactic herpes simplex virus type 2 vaccine adjuvanted with a universal CD4 T cell helper peptide induces long-term protective immunity against lethal challenge in mice. Int Immunopharmacol 61:100–108. https://doi.org/10.1016/j.intimp.2018.05.024

Madeira F, Pearce M, Tivey AR, Basutkar P, Lee J, Edbali O, Madhusoodanan N, Kolesnikov A, Lopez R (2022) Search and sequence analysis tools services from EMBL-EBI in 2022. Nucleic Acids Res 50:276–279. https://doi.org/10.1093/nar/gkac240

Majee P, Jain N, Kumar A (2021) Designing of a multi-epitope vaccine candidate against Nipah virus by in silico approach: a putative prophylactic solution for the deadly virus. J Biomol Struct Dyn 39(4):1461–1480. https://doi.org/10.1080/07391102.2020.1734088

Nayak AK, Chakraborty A, Shukla S, Kumar N, Samanta S (2024) An immunoinformatic approach for developing a multi-epitope subunit vaccine against Monkeypox virus. Silico Pharmacol 12(1):42. https://doi.org/10.1007/s40203-024-00220-5

Nguyen TL, Kim H (2024) Immunoinformatics and computational approaches driven designing a novel vaccine candidate against Powassan virus. Sci Rep 14(1):5999. https://doi.org/10.1038/s41598-024-56554-9

Peterson EE, Barry KC (2021) The natural killer–dendritic cell immune axis in anti-cancer immunity and immunotherapy. Front Immunol 11:621254. https://doi.org/10.3389/fimmu.2020.621254

Rapin N, Lund O, Bernaschi M, Castiglione F (2010) Computational immunology meets bioinformatics: the use of prediction tools for molecular binding in the simulation of the immune system. PLoS ONE 5(4):9862. https://doi.org/10.1371/journal.pone.0009862

Reynisson B, Barra C, Kaabinejadian S, Hildebrand WH, Peters B, Nielsen M (2020) Improved prediction of MHC II antigen presentation through integration and motif deconvolution of mass spectrometry MHC eluted ligand data. J Proteome Res 19(6):2304–2315. https://doi.org/10.1021/acs.jproteome.9b00874

Robinson CL (2018) Advisory committee on immunization practices recommended immunization schedule for children and adolescents aged 18 years or younger—United States, 2018. MMWR 67

Sartorius B, VanderHeide JD, Yang M, Goosmann EA, Hon J, Haeuser E, Cork MA, Perkins S, Jahagirdar D, Schaeffer LE, Serfes AL (2021) Subnational mapping of HIV incidence and mortality among individuals aged 15–49 years in sub-saharan Africa, 2000–18: a modelling study. Lancet HIV 8(6):363–375. https://doi.org/10.1016/S2352-3018(21)00051-5

Satterfield BA, Cross RW, Fenton KA, Agans KN, Basler CF, Geisbert TW, Mire CE (2015) The immunomodulating V and W proteins of Nipah virus determine disease course. Nat Commun 6(1):7483. https://doi.org/10.1038/ncomms8483

Sharma SK, Srivastava S, Kumar A, Srivastava V (2021) Anticipation of antigenic sites for the goal of Vaccine Designing against Nipah Virus: an Immunoinformatics Inquisitive Quest. Int J Pept Res Ther 27(3):1899–1911. https://doi.org/10.1007/s10989-021-10219-7

Silva-Arrieta S, Goulder PJ, Brander C (2020) In silico veritas? Potential limitations for SARS-CoV-2 vaccine development based on T-cell epitope prediction. PLoS Pathog 16(6):1008607. https://doi.org/10.1371/journal.ppat.1008607

Srivastava S, Verma S, Kamthania M, Saxena AK, Pandey KC, Pande V, Kolbe M (2023) Exploring the structural basis to develop efficient multi-epitope vaccines displaying interaction with HLA and TAP and TLR3 molecules to prevent NIPAH infection, a global threat to human health. PLoS ONE 18:0282580. https://doi.org/10.1371/journal.pone.0282580

Venkatarajan MS, Braun W (2001) New quantitative descriptors of amino acids based on multidimensional scaling of a large number of physical–chemical properties. Mol Model Ann 7:445–453. https://doi.org/10.1007/s00894-001-0058-5

Wold S, Jonsson J, Sjörström M, Sandberg M, Rännar S (1993) DNA and peptide sequences and chemical processes multivariately modelled by principal component analysis and partial least-squares projections to latent structures. Anal Chim Acta 277:239–253. https://doi.org/10.1016/0003-2670(93)80437-P

Yuan S, Chan HS, Hu Z (2017) Using PyMOL as a platform for computational drug design. WIREs Comput Mol Sci 7(2):1298. https://doi.org/10.1002/wcms.1298

Zeb S, Mushtaq M, Ahmad M, Saleem W, Rabaan AA, Naqvi BSZ, Garout M, Aljeldah M, Al Shammari BR, Al Faraj NJ, Al-Zaki NA (2022) Self-medication as an important risk factor for antibiotic resistance: a multi-institutional survey among students. Antibiotics 11(7):842. https://doi.org/10.3390/antibiotics11070842

Zhang Y, Huo F, Cao Q, Jia R, Huang Q, Wang ZA, Theodorescu D, Lv Q, Li P, Yan C (2022) FimH confers mannose-targeting ability to Bacillus Calmette-Guerin for improved immunotherapy in bladder cancer. J ImmunoTher Cancer 10(3). https://doi.org/10.1136/jitc-2021-003939

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.