



# Anticipation of Antigenic Sites for the Goal of Vaccine Designing Against Nipah Virus: An Immunoinformatics Inquisitive Quest

Suraj kumar Sharma<sup>1</sup> · Shivani Srivastava<sup>1</sup> · Ajay Kumar<sup>1</sup> · Vivek Srivastava<sup>1</sup>

Received: 12 January 2021 / Accepted: 12 April 2021  
© The Author(s), under exclusive licence to Springer Nature B.V. 2021

## Abstract

With time, the Nipah virus has been proved as a fatal and dangerous pathogen for humanity. Nipah virus has its origin from bats and severely affects the respiratory as well as neurological organs. Regular outbreaks and unavailability of proper treatment for Nipah virus infection, demands the designing of vaccine for this disease. This prediction study was conducted to explore B cell epitopes from the Nipah virus's proteome using the immunoinformatics approach. In this curious quest of anticipation of antigenic sites for the *Insilico* peptide vaccine for the Nipah virus, nine NV-B strain proteins were retrieved for further series of investigations. After sequential refining through immunoinformatics approaches, a total of 26 epitopes was selected to perform molecular modeling and docking. PEPstrMOD and Swiss model, respectively performed 3D modeling of epitopes with their respective alleles. Based on minimum binding energy, four epitopes viz. LHLGNFVRR, LNLSPLIQR, YHNMSPINR and FRRNNAIAF were predicted as promiscuous B cell epitopes. Based on low binding affinity and high population coverage worldwide, epitope LHLGNFVRR was finally selected. Increased Stability of the LHLGNFVRR- HLA DRB\_1301 complex during simulation studies exhibit it as the most promising vaccine bidder. So complex of LHLGNFVRR-HLA DRB\_1301 has shown most significance result for vaccine and for further validation and confirmation, wet lab and clinical trials can provide the potential of predicted peptides for the subunit vaccine.

**Keywords** Epitopes · Nipah virus · Docking · Immunoinformatics · Simulation · MHC Class

## Introduction

Nipah virus infection is an emerging disease driven by bats. Malaysia was the country in which the first case of the virus was reported. Since then, it has affected different regions of the world, especially Asia (Aditi and Shariff 2019). The term Nipah has been derived from the name of the village Sungai Nipah in Malaysia (CDC 2017). Infection caused by the Nipah virus has become one of the looming zoonotic diseases as it can be transmitted from animals to humans. Nipah virus belongs to the subfamily Paramyxoviridae. Phosphoprotein, fusion, matrix, glycoprotein, nucleoprotein, and RNA polymerase are the six genes that make up the virus's whole genome. It also consists of three non-structural proteins that depend on the host's innate immune response in vitro (King et al. 2011). Fruit bats are the primary host

of this life-threatening pathogen. These fruit bats mainly feed on the fruits and nectar and thus limit the transmission area of the virus. Asia, Australia, East Africa are the areas primarily confined with the bats and therefore are more prone to frequent outbreaks of the Nipah virus (Epstein et al. 2006). It is a highly contagious disease and can transmit in the community through close association with diseased individuals or animals (WHO 2019). Symptoms of Nipah virus infection include fever, headache, muscle pain, sore throat, and vomiting. In severe cases, these symptoms are followed by neurological disorders and encephalitis. After 24 to 48 h, an infection may be followed by a coma. The incubation period ranges from 4 to 14 days (WHO 2018).

In Malaysia, states viz., Perak, and Negri Sembilan witnessed the Nipah virus cases for the first time in the year 1998. Few cases were also reported from Bukit Pelandok. These outbreaks reported a 39.6% mortality rate (Aljofan 2013). In 2001, the Nipah virus outbreak occurred in Meherpur, Bangladesh (Hsu et al. 2005; Hossain et al. 2008), and Siliguri, India (Chadha et al. 2006). Years 2003, 2004, and 2005 both in India and Bangladesh witnessed

✉ Vivek Srivastava  
viveksrivastavabio@gmail.com

<sup>1</sup> Department of Biotechnology, Rama University Uttar Pradesh, Kanpur 209217, India

regular outbreaks (Chadha et al. 2006, WHO 2018). In 2018, 19 Nipah Virus cases were reported in Kozhikode and Malappuram districts in Kerala State. Previously in 2001 and 2007 also, India had witnessed two outbreaks in West Bengal. Invasion of this virus in host cells takes place with the help of cell to cell fusion. Nipah virus glycoprotein and fusion protein play a significant role in cell to cell spread of the virus. After binding the receptor, glycoprotein performs conformational changes in the fusion protein, leading to viral entry (Wong et al. 2002, Liu et al. 2015). A wide range of methods are available for the early detection of the Nipah virus's presence in the host, viz. molecular assays, Reverse Transcriptase-Polymerase chain reaction (RT-PCR), Nested PCR, Next-generation sequencing, Enzyme-linked immunosorbent assay (ELISA), etc. Among all these methods, RT-PCR is most commonly used because of its extreme sensitivity and accuracy (Mazzola and Kelly-Cirino 2019). Several in vitro researches have proved the secretion of Interferon- $\beta$  (IFN- $\beta$ ), antibodies like IgM and IgG from infected cells (Lo et al. 2010, Chong et al. 2001). Many animal models have been en route to understand the mechanism of pathogenicity and transmission of the Nipah virus into the host. The study of the virus in animal models had contributed to the development of insilico vaccines. African green monkeys, Hamsters, pigs, and Ferrets are the most successful animal models for investigating the Nipah virus (De Wit and Munster 2014). Phosphoprotein and V protein of Nipah virus share an intrinsic N-terminal domain, which has been determined by NMR spectroscopy. This intrinsically disordered domain having 1–104 amino acids plays an essential role in viral genome replication (Schiavina et al. 2020).

According to the Centre For Disease Control and Prevention report, the Nipah virus is an emerging infectious disease of public health importance. Few drugs and monoclonal antibody viz., Remdesivir and m102.4 had shown excellent therapeutic potential. It has been proved experimentally in nonhuman primates. Using antivirals as a prerequisite in therapeutics has been revolutionized greatly for infection against Nipah Virus (Lo et al. 2019). There is no remedy, and it kills about 75% of people it infects (CDC 2019). Nipah Virus has dawned as one of the most dangerous and fatal zoonotic infections throughout the world. The ever-expanding human population, widening interaction between humans and animals, dynamic environmental conditions, and poor sanitation are the main reasons for the outbreaks of zoonotic diseases (Kazi et al. 2018). One of the most protrusive ways to deal with this challenge is vaccination. This terminology was initially used by Edward Jenner, popularly known as the father of vaccinology, 1796. With time, the vaccine has proved to be the most potent approach to combat such dreadful diseases. Computational methods had created a boom in the finding of epitopes and, thus, the development of vaccines. Here we present a primarily

centralized study on exploring the Nipah virus proteome using an immunoinformatics approach to predict potential B cell epitopes for vaccine designing. In this study, all the Nipah virus proteins were collected from the UniProtKB database and were inspected for allergenicity. Further B cell epitopes were identified, and their potency was checked by using the VaxiJen server. A non-toxicity test was also performed on potential epitopes. This was followed by tertiary structure modeling of selected epitopes and their respective alleles. The effectiveness of these peptides with their corresponding alleles was examined by performing molecular docking and population coverage analysis. The strength of the peptide complex with its respective allele was substantiated after conducting simulation studies. This course of deciphering of potential vaccines bidders had fired up the processes required for epitope screening. It focused on the most crucial vaccine and antibody design and development while saving time, minimizing cost, and most notably, eliciting a proper immune response (Kulkarni et al. 2013).

## Methodology

### Protein Sequence Retrieval

The Nipah virus NV-B strain (GenBank accession number AY988601) was retrieved from the UniProtKB database. The genome of the NV-B strain is 18252 nucleotides in length (Harcourt et al. 2004). NV-B strain of Nipah virus has an immensely shorter incubation period (Kulkarni et al. 2013; Rahman et al. 2012), has a high mortality rate, and is exceptionally lethal comparison to other strains (Clayton et al. 2012; Sazzad et al. 2013). In most cases, infection with this strain reflects encephalitis and respiratory system problems (Goh et al. 2000; Chadha et al. 2006). This strain has a high propensity for a human to human transmission (Griffin et al. 2019). All these factors have contributed to the selection of the strain for the interesting study. UniProtKB is a freely accessible database that contains a large amount of information on proteins derived from the research literature. In this study, a total of 9 proteins viz. Polymerase, Nucleocapsid, Fusion, W, Glycoprotein, Phosphoprotein, Matrix, V, and C, were retrieved. Sequences of these proteins were saved in the FASTA format for the subsequent sequential investigation.

### Allergenic Protein Prediction

As proteins with allergic properties cannot be considered and thus cannot be proved to be a capable vaccine contender. Therefore, it is necessary to separate non-allergens and allergens from retrieved protein sequences. To achieve this objective, the AllgPred server was used (Saha and Raghava

2006). The tool anticipates allergic properties of the proteins by using six peculiar ways. This research has used the Support- vector machines (SVM) method based on amino acid composition.  $-0.4$  was the cut-off value. 88.87% was the sensitivity, and 81.86% can be the precision ascertained at  $-0.4$  threshold value. All non-allergens were selected for the subsequent investigation.

## B Cell Epitope Prediction

The B cell epitopes are peptide fragments that can provoke a proper immune response and are extremely important for vaccine design utility from the pathogen's proteome. Due to B cell epitopes' importance, NETMHCII 2.3 server was used to predict B cell epitopes derived from protein sequences of the Nipah virus (Jensen et al. 2018). It gives information about the binding of predicted epitopes to Human Leukocyte Antigen (HLA) Alleles viz. HLA-DR, HLA-DQ, HLA-DP, and mouse class II alleles using artificial neuron networks. In this study, we have identified the epitope for the allele DRB1\_0101, DRB1\_0701, DRB1\_1301. Peptide length was chosen nine, and the threshold value was set at  $-99.9$ . The potential peptides were selected for further analysis by the VaxiJen server.

VaxiJen is the alignment independent prediction server that allows antigen classification based on physiochemical properties. Prediction accuracy ranges from 70 to 89% (Doytchinova and Flower 2007). 1.1 was selected as the cut-off value, and epitopes with a score greater than 1.1 can tether with significant histocompatibility complex-II (MHC II) molecules with high-level confidence and thus selected for further analysis.

## Toxicity Prediction of the Epitopes

Toxin Pred is an in silico method used to predict the non-toxicity of selected epitopes (Gupta et al. 2013). The anticipation of this server is based on quantitative matrix-based position-specific scores. It also gives details of different physicochemical properties of epitopes like molecular weight, hydrophobicity, and hydrophilicity.

## Molecular Modeling of Epitopes and HLA Alleles

After analyzing the allergenicity, toxicity, and anticipation of epitopes, modeling of tertiary structures of the selected epitopes was performed by the PEPstrMOD server (Singh et al. 2015). This server forebodes the tertiary structure of peptides having 7 to 25 residues. It also deals with modifications like acetylation, cyclization, post-translational, etc. In this study, non-toxic epitopes were modeled using it and were saved in the Protein data bank format (PDB) format for further analysis.

The Swiss Model server was used to generate the 3D structure of HLA alleles that will bind to selected epitopes (Waterhouse et al. 2018). It is a web-based integrated service that guides the user for homology modeling of different complex proteins. Here we have performed the modeling of HLA alleles viz. DRB1\_0101, DRB1\_0701, and DRB1\_1301, by using it. Depending on the complexity of modeling, this server provides three modes which are as follows: automated mode, alignment mode, and project mode. In this research, an alignment mode was chosen to generate 3D structures of the alleles mentioned above.

## Molecular Docking of Epitopes and HLA Alleles

To know the interaction between selected epitopes and their corresponding alleles, molecular docking was performed through AutoDock Vina software (Trott and Olson 2010). Before performing docking, certain modifications need to be done both in the ligand and the receptor. These modifications were performed using AutoDock MGL tools.

To perform *Insilico* molecular docking, HLA-DRB1\_0101, HLA-DRB1\_0701, HLA-DRB1\_1301 were selected as receptors, and the crystal structure of these receptors viz. 4AH2, 3C5J, and 6CQL were retrieved using the Research Collaboratory for Structural Bioinformatics (RCSB) protein data bank. To prepare receptors for docking, removed water molecules as they will cause hindrance during docking studies. Then, polar hydrogen and Kollman charges were added, and finally, the receptor was saved in pdbqt format. All 26 ligands were also prepared by using AutoDock MGL tools and were held in pdbqt form.

The grid box was set up at 40, 40, 40, and the docking was completed at  $0\text{\AA}$  spacing parameter for identifying binding affinity. Protein–ligand interaction was visualized by using the PyMOL visualization tool. Based on low binding energy, four epitopes were selected for final refining. The more negative binding affinity means more vital interaction between epitope and allele complex.

## Binding Affinity Prediction and Population Coverage Analysis

The binding affinity of epitopes with MHCII molecules was predicted with the MHCpred server, which uses the additive method to envisioned the binding affinity of anticipated nano peptides with MHC class II alleles (Guan et al. 2003). In this research, the amino acid contribution model was chosen, and peptides showing a binding affinity of  $<500$  nM were selected for further analysis.

MHC alleles show great diversity among different ethnicities and geographical areas. Therefore, during vaccine designing, population coverage must be considered as an essential criterion. Population coverage analysis was

performed using the Immune Epitope Database (IEDB) Population Coverage analysis tool (Bui et al. 2006).

## Dynamic Simulation Studies

Molecular dynamics (MD) and Simulation of the peptide—allele association were studied using the MDWeb tool (Hospital et al. 2012). It is a web interface to study molecular dynamics trajectories and to perform the simulation. It is essential to perform a simulation to determine the stability of the epitope -allele complex. A standard MD setup was used to equilibrate the system. The simulation was run for ten nanoseconds, and the output frequency was set at 500 steps. The distance between Alpha carbon atoms was 3.8 Å and was analyzed for trajectory. Output was given in Root Mean Square Deviation (RMSD) values and B factor values to validate the epitope-allele complex's stability.

## Results

Recent years had witnessed an emanation of zoonotic diseases like the Nipah virus in various parts of the world. With time, the Nipah virus and its associated outbreaks had created a menace for both humanity and society. Despite the high fatality rates of this disease, there is no treatment or vaccine available until now. Present research work deals with the exploration and characterization of epitopes from the proteome of the Nipah virus. These epitopes have a high possibility of eliciting proper and robust immune responses. Therefore, the description of B cell epitopes screened from the NV-B strain is of utmost importance. First and foremost, the step in *insilico* work of vaccine designing is the selection of non-allergen nine protein. The most promising vaccine candidates must be non-allergen. It must not cause any allergy in the host. The elimination of non-allergens from allergens was performed by the AlgPred tool. After retrieving all proteins of strain NV-B from UniProtKB, screening

and filtration were performed to obtain non-allergic proteins. Of all the 9 proteins encoded by the virus, 4 were non-allergens (Table 1). All nine proteins and their protein ID and the AlgPred Prediction score results have been summarized in Table 1.

Following this step, non-allergens were subjected to NETMHC II 2.3 server to predict B cell epitopes. These epitopes were obtained at threshold value – 99.9 and epitopes obtained were nano peptides. Besides this, predicted epitopes were also analyzed by the VaxiJen tool. For screening through VaxiJen server 1.1 was taken as cut off value. A Total of 26 epitopes could bind with MHC II molecules with high stability i.e., having a VaxiJen score  $\geq$  of 1.1 was obtained. Among these 26 epitopes, four epitopes were found to bind with MHC allele DRB1\_0101, three epitopes were found to bind with MHC allele DRB1\_0701, and 19 epitopes were found to bind with MHC allele DRB1\_1301 (Table 2). All these epitopes, along with their corresponding alleles and VaxiJen score, have been summarized in Table 2.

Toxic peptides can never be embodied as vaccine binders. Therefore, one of the crucial steps in vaccine designing is determining the non-toxicity of selected epitopes. Toxin Pred performed this investigation. It was found that all 26 epitopes were nontoxic (Table 3). The result of the analysis, along with the SVM score, molecular weight, hydrophobicity, and hydrophilicity, has been summarized in Table 3.

The Swiss model performed the generation of the 3D structure of HLA alleles. It generated the 3D or tertiary framework of HLA alleles: DRB1\_0101, DRB1\_0701, and DRB1\_1301 were performed by the Swiss model (Fig. 1). To generate tertiary structures of these alleles' proteins with PDB ID 4AH2, 3C5J and 6CQL were used as templates for alleles DRB1\_0101, DRB1\_0701, and DRB1\_1301, respectively. 3D models of the alleles were visualized by the PyMOL visualization tool. PEPstrMOD performed the modeling of selected epitopes. Tertiary structures of alleles have been represented in Fig. 1.

**Table 1** Anticipation of allergic properties of proteins along with their names, protein Id, and AlgPred score

S. No	Protein name	Protein Id	Score	AlgPred prediction
<b>1</b>	<b>RNA-directed RNA polymerase L</b>	Q4VCP4	– 0.8954299	<b>Non-allergen</b>
<b>2</b>	<b>Nucleocapsid protein</b>	Q4VCQ2	– 1.6635	<b>Non-allergen</b>
3	Fusion protein	Q4VCP6	– 0.0698	Allergen
4	Glycoprotein	Q4VCP5	0.42194	Potential allergen
5	W protein	Q4VCP9	– 0.3831	Allergen
6	V protein	Q4VCQ0	– 0.2837	Allergen
<b>7</b>	<b>Matrix protein</b>	Q4VCP7	– 1.0104617	<b>Non- allergen</b>
8	Phosphoprotein	Q4VCQ1	– 0.0534	Allergen
<b>9</b>	<b>C protein</b>	Q4VCP8	– 1.0820195	<b>Non-allergen</b>

Selected proteins are represented in bold

**Table 2** List of B cell epitopes as anticipated by NETMHCII 2.3 server and the result of VaxiJen analysis indicating antigenicity of epitopes

Protein Id	Allele	Epitopes	Binding affinity (nM)	VaxiJen score	Antigen/non-antigen
AAY43917.1	DRB1_0101	FNYLNLSRI	598.8	1.1049	Antigen
	DRB1_0101	YHNMSPNIR	1095.6	1.3495	Antigen
	DRB1_0701	FKDTKSSEL	98	1.1432	Antigen
	DRB1_0701	ILDFRSKLM	221.3	1.5359	Antigen
	DRB1_1301	MSPNIRNRY	14.2	2.0730	Antigen
	DRB1_1301	WAFEIHHRR	31.6	1.5661	Antigen
	DRB1_1301	IQPRLVSRL	37.1	1.4218	Antigen
	DRB1_1301	LLCLQKTVK	44.4	1.2822	Antigen
	DRB1_1301	LRMNLRALG	44.6	1.3042	Antigen
	DRB1_1301	RLRMNLRAL	46.7	1.2762	Antigen
	DRB1_1301	LNLSPLIQR	50.4	1.4413	Antigen
	DRB1_1301	YASLTYLRR	56.6	1.1006	Antigen
	DRB1_1301	VLILDFRSK	60.6	2.1075	Antigen
	DRB1_1301	VIFKIKNSQ	65.5	1.1178	Antigen
	DRB1_1301	FVYSKKIHY	65.6	1.2617	Antigen
	DRB1_1301	FAKMTYKMR	68.7	1.3595	Antigen
	DRB1_0101	LLNLSRLA	53.9	1.4762	Antigen
	DRB1_1301	LNLRSRLAA	15.1	1.7065	Antigen
	DRB1_1301	ATLTTKIRI	27.3	1.2084	Antigen
	DRB1_1301	LRSRLAAKA	28	1.2209	Antigen
AAY43911.1	DRB1_1301	IRFGLETRY	30.5	2.2023	Antigen
	DRB1_1301	SLLNLSRL	37.1	1.6202	Antigen
	DRB1_1301	LGQKSARHH	57.8	1.4277	Antigen
AAY43914.1	DRB1_0101	FFLSITKLN	1022.8	1.1684	Antigen
	DRB1_0701	FRRNNAIAF	118.9	1.1890	Antigen
	DRB1_1301	LHLGNFVRR	10.4	1.3169	Antigen

Cut off value for analysis by VaxiJen server  $\geq 1.1$ 

Molecular docking between nontoxic, non-allergic antigenic peptides and their corresponding alleles was performed by AutoDock Vina software. The binding pattern of all 26 epitope-allele complexes are depicted in the form of binding energies in terms of kcal/mol (Table 4).

Epitopes viz., YHNMSPNIR, LNLSPLIQR, FRRNNAIAF, and LHLGNFVRR showed binding energies of  $-7.4$  kcal/mol,  $-6.8$  kcal/mol,  $-6.8$  kcal/mol, and  $-6.9$  kcal/mol, respectively. A stable complex of YHNMSPNIR- 4AH2 forms 3 Hydrogen bonds (Fig. 2).

The stable complex of LNLSPLIQR-6CQL and FRRNNAIAF- 3C5J forms 5 and 2 Hydrogen bonds, respectively (Figs. 3 and 4). Similarly, the LHLGNFVRR-6CQL complex structures ten hydrogen bonds (Fig. 5).

The binding affinity of selected epitopes with different alleles (HLA-DRB1\_0101, HLA-DRB1\_0401, HLA DRB1\_0701) was determined with MHC Pred. Peptides showing binding affinity of  $IC_{50} < 500$  nM are contemplated as excellent binders. LHLGNFVRR was found to bind with HLA- DRB1\_0101 and HLA- DRB1\_0701.

LNLSPLIQR was found to bind with HLA- DRB1\_0101 and HLA- DRB1\_0401. YHNMSPNIR was found to bind with HLA DRB1\_0101. FRRNNAIAF was found to bind with HLA-DRB1\_0101, HLA- DRB1\_0701, HLA DRB1\_0401 (Table 5). The results of the prediction of binding affinity are summarized in Table 5.

Population coverage analysis was performed by the IEDB population coverage tool. On investigation, it was found that the worldwide highest coverage percentage (38.05%) was obtained by peptide FRRNNAIAF (Fig. 6). Both peptides, viz., LHLGNFVRR, and LNLSPLIQR, show 22.06% population coverage, respectively (Fig. 7). The least population coverage was covered by YHNMSPNIR (11.53%) (Fig. 8). The result of population coverage analysis has been summarized in Figs. 6, 7 and 8.

Among the four epitopes, the final epitope was selected based on two factors. One is high population coverage, and another one is low binding affinity. YHNMSPNIR had the lowest binding energy ( $-7.4$ ), but it was not selected as the final vaccine candidate because of its meager worldwide



**Table 3** Result showing Investigation of toxic nature of antigenic peptides, their molecular weight, hydrophobicity, and hydrophilicity

Peptides	SVM Score	Toxin/non-toxin	Molecular weight	Hydrophobicity	Hydrophilicity
FNLYLNLSRI	− 1.06	Non-toxin	1139.4	− 0.1	− 0.72
YHNMSPNIR	− 0.72	Non-toxin	1131.4	− 0.31	− 0.24
FKDTKSSEL	− 1.29	Non-toxin	1054.2	− 0.34	0.88
ILDFRSKLM	− 1.37	Non-toxin	1122.5	− 0.13	0.01
MSPNIRNRY	− 0.37	Non-toxin	1150.44	− 0.46	0.14
WAFEIHHRR	− 1.19	Non-toxin	1251.54	− 0.33	− 0.02
IQPRLVSRL	− 1.3	Non-toxin	1081.46	− 0.25	− 0.04
LLCLQKTVK	− 0.41	Non-toxin	1045.49	− 0.1	− 0.23
LRMNLRALG	− 1.07	Non-toxin	1043.43	− 0.21	− 0.11
RLRMNLRAL	− 1.07	Non-toxin	1142.56	− 0.42	0.22
LNLSPLIQR	− 1.18	Non-toxin	1053.41	− 0.12	− 0.39
YASLTYLRR	− 1.27	Non-toxin	1142.44	− 0.29	− 0.31
VLILDFRSK	− 1.49	Non-toxin	1090.46	− 0.1	− 0.01
VIFKIKNSQ	− 1.05	Non-toxin	1076.44	− 0.13	− 0.1
FVYSKKIHY	− 1.3	Non-toxin	1184.53	− 0.1	− 0.51
FAKMTYKMR	− 0.57	Non-toxin	1175.59	− 0.3	0.08
LLNLSRLA	− 1	Non-toxin	1055.42	− 0.23	− 0.13
LNLRSRLAA	− 1.09	Non-toxin	1013.33	− 0.26	0.01
ATLTTKIRI	− 1.32	Non-toxin	1016.38	− 0.13	− 0.12
LRSRLAAKA	− 0.65	Non-toxin	985.31	− 0.34	0.47
IRFGLETRY	− 0.5	Non-toxin	1154.46	− 0.25	0.02
SLLNLSRSL	− 0.85	Non-toxin	1071.42	− 0.28	− 0.04
LGQKSARHH	− 1.09	Non-toxin	1033.29	− 0.41	0.36
FFLSITKLN	− 1.51	Non-toxin	1082.44	0.09	− 0.81
FRRNNAIAF	− 0.97	Non-toxin	1108.38	− 0.26	− 0.16
LHLGNFVRR	− 1.33	Non-toxin	1111.45	− 0.24	− 0.21

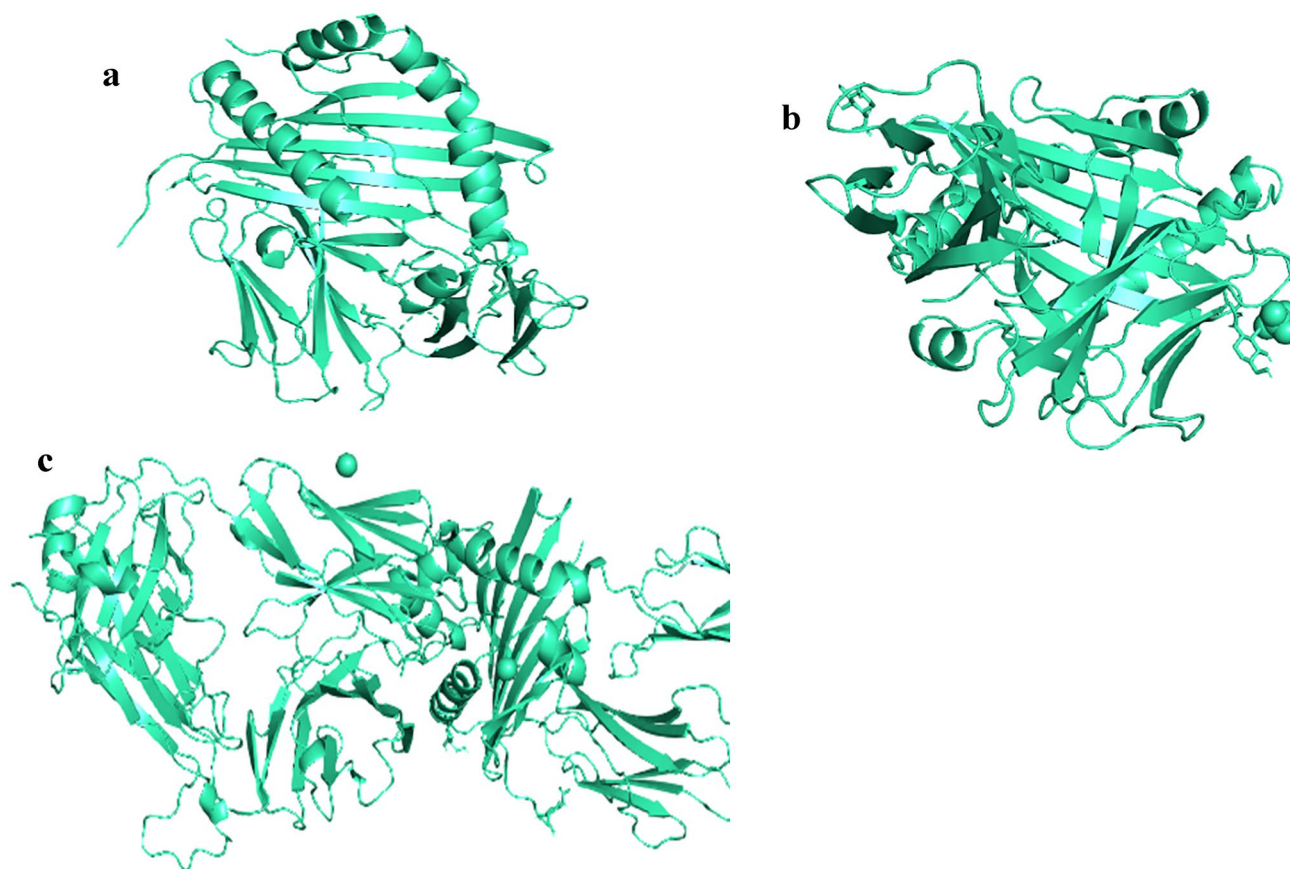
population coverage. Epitope LHLGNFVRR had a low binding affinity and good population coverage; therefore, it was chosen as a final vaccine candidate. The complex of LHLGNFVRR- 6CQL exhibits ten hydrogen bonds. To analyze the study of stability and interaction of epitope—allele complex, a simulation dynamics study was performed on the final vaccine candidate, i.e., LHLGNFVRR. The MDWeb simulation tool's output was given in terms of RMSD values and B factor values to validate protein–ligand interaction more accurately. Docked complex shows best RMSD values between 0.1 and 1.0 Å (Fig. 9) and B factor scores between 1 and 25 Å<sup>2</sup> (Fig. 10). It offers the stability of the final epitope-allele complex. RMSD values and B factor values are depicted in Figs. 9 and 10, respectively.

## Discussion

In silico vaccine designing, also known as reverse vaccinology, is a domain that comprises epitope screening, selection, and mapping for vaccine designing, using computational tools and software. These in silico exploration methods had increased the process of novel vaccine preparation against

emerging biothreats (De Groot et al. 2003). Wet lab experiments in the field of immunology are costly. These experiments also generate a large amount of data which is needed to be analyzed with intensive accuracy and precision (Manzoni et al. 2018). To deal with such challenges, bioinformatics tools had emerged drastically. Using computational methods had reduced both costs and time, especially in vaccine formulation for emerging infectious diseases (Peng et al. 2019). Immunoinformatics-based subunit vaccines encompass more safety and better potency to provoke an immune response by predicting epitopes (Zhou et al. 2009). The development of vaccines with the help of the immunoinformatics approach has been fabricated against many diseases like cholera (Nezafat et al. 2016), cancer (Zhang 2017), and helicobacter (Meza et al. 2017). This present study deals with the proteomic screening of the Nipah virus and using a reverse vaccinology approach to explore B cell epitopes that can provoke a better immune response in humans.

In this research work, the bioinformatics tool viz. Vaxi-Jen, NETMHC II2.3, Toxin Pred, MD Web, and AutoDock vina were used to determine the best vaccine bidder. In the first step of computational vaccinology, a total of 9 proteins of the NV-B strain of the Nipah virus were retrieved from



**Fig. 1** 3D modeled structure of HLA Class II alleles **a** Molecular structure of HLA DRB1\_0101, **b** Molecular structure of HLA DRB1\_0701, **c** Molecular structure of HLA DRB1\_1301

the UniProtKB database. These proteins were screened and filtered to find out epitopes by using NETMHC II 2.3 and VaxiJen server. Only those 26 epitopes were selected for further analysis to bind with its corresponding allele with high stability. All 26 epitopes were analyzed for toxicity and other properties like molecular weight, hydrophobicity, and hydrophilicity. 3D modeling of epitopes, as well as alleles, was generated by the PEPstrMOD and Swiss model, respectively. To create 3D modeling of alleles, proteins with PDB ID 4AH2, 3C5J, and 6CQL were used as templates for alleles DRB1\_0101, DRB1\_0701, DRB1\_1301, respectively. To select the best epitope among these, molecular docking was accomplished by AutoDock vina. Based on low binding affinity, four epitopes viz, YHNMSPNIR (−7.4 kcal/mol), LNLSPLIQR (−6.8 kcal/mol), FRRNNAIAF (−6.8 kcal/mol), and LHLGNFVRR (−6.9 kcal/mol). The presence of hydrogen bonds depicts more clearly these epitope-allele complexes viz., YHNMSPNIR-4AH2, LNLSPLIQR-6CQL, FRRNNAIAF-3C5J, and LHLGNFVRR-6CQL. To predict the binding affinity of these 4 epitopes with different alleles viz. HLA-DRB1\_0101, DRB1\_0401, and DRB1\_1301, MHC Pred server was used.

The binding affinity of predicted epitopes with alleles was selected based on  $IC_{50}$  value. Epitopes having  $IC_{50}$  values less than 500 nM are featured as suitable binders. LHLGNFVRR was found to bind with HLA-DRB1\_0101 and HLA-DRB1\_0701. LNLSPLIQR was found to bind with HLA-DRB1\_0101 and HLA-DRB1\_0401. YHNMSPNIR was found to bind with HLA-DRB1\_0101. FRRNNAIAF was found to bind with HLA-DRB1\_0101, HLA-DRB1\_0701, HLA-DRB1\_0401. Population coverage analysis was also performed on 4 epitopes using the IEDB population coverage tool. Based on low binding energy and satisfactorily population coverage, epitope LHLGNFVRR was finally chosen to perform dynamic simulation studies. Simulation performed on LHLGNFVRR- HLA-DRB1\_1301 complex was completed, which validated the stability of the complex. The result of the simulation was given in the form of RMSD and B Factor values. This prediction study had anticipated an epitope LHLGNFVRR which can provoke an excellent immune response. This epitope can bind with its corresponding allele with superior stability, is non-toxic, and low binding energy. All these properties confer that the

**Table 4** List showing Binding affinity between 26 selected epitopes and their corresponding alleles as predicted by AutoDock Vina software

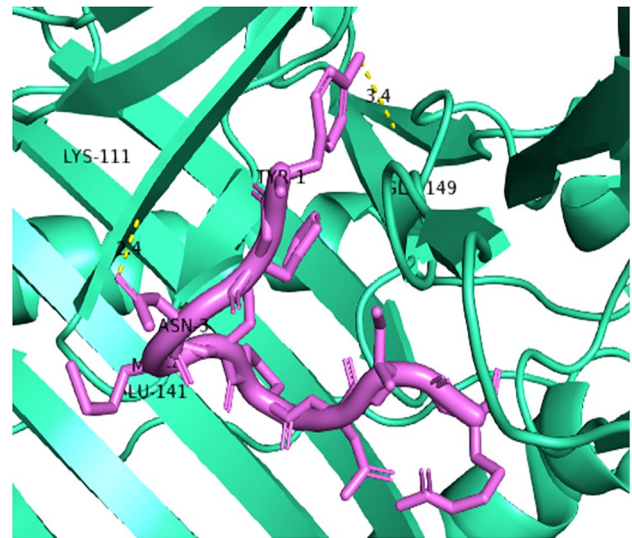
S. No	Epitopes	Allele	Binding affinity (kcal/mol)
1	FNYLNLSRI	DRB1_0101	− 6.7
2	<b>YHNMSPNIR</b>	<b>DRB1_0101</b>	<b>− 7.4</b>
3	FKDTKSSEL	DRB1_0701	− 6.4
4	ILDFRSKLM	DRB1_0701	− 6.5
5	MSPNIRNRY	DRB1_1301	− 5.7
6	WAFEIHRR	DRB1_1301	− 6.1
7	IQPRLVSRL	DRB1_1301	− 6.3
8	LLCLQKTVK	DRB1_1301	− 6
9	LRMNLRALG	DRB1_1301	− 6.2
10	RLRMNLRAL	DRB1_1301	− 5.5
11	<b>LNLSPLIQR</b>	<b>DRB1_1301</b>	<b>− 6.8</b>
12	YASLTYLRR	DRB1_1301	− 6.5
13	VLILDFRSK	DRB1_1301	− 5.3
14	VIFKIKNSQ	DRB1_1301	− 5.6
15	FVYSKKIHY	DRB1_1301	− 6.4
16	FAKMTYKMR	DRB1_1301	− 5.9
17	LLNLSRLA	DRB1_0101	− 5.5
18	LNLSRLAA	DRB1_1301	− 5.7
19	ATLTTKIRI	DRB1_1301	− 6.3
20	LSRLAAKA	DRB1_1301	− 5.6
21	IRFGLETRY	DRB1_1301	− 6.2
22	SLNLSRL	DRB1_1301	− 5.6
23	LGQKSARHH	DRB1_1301	− 6
24	FFLSITKLN	DRB1_0101	− 6.7
25	<b>FRRNNAIAF</b>	<b>DRB1_0701</b>	<b>− 6.8</b>
26	<b>LHLGNFVRR</b>	<b>DRB1_1301</b>	<b>− 6.9</b>

Final selected epitopes are represented in bold

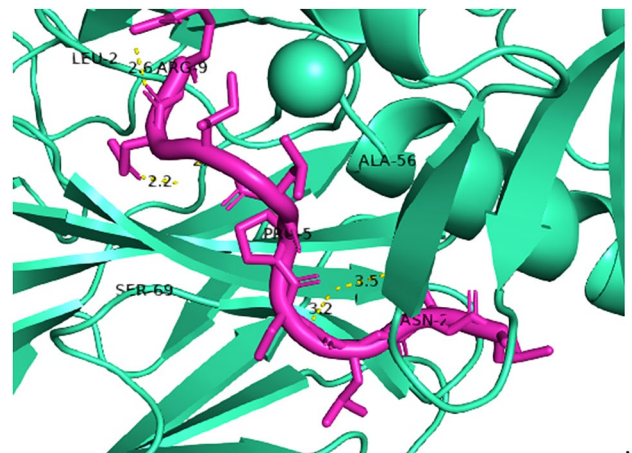
selected peptide will be proved as the best vaccine candidate and further validated by clinical trials.

## Conclusion

Time and again, World has faced outbreaks of several life-threatening diseases. Although there have been advancements in technology to combat and fight some significant infections, we still lack some zoonotic diseases like the Nipah virus. This virus is a potent threat to humanity and society. Vaccines have always been proved as the greatest weapon in the battle between humans and diseases. While it may have contributed to eradicating deadly diseases, there are still a lot of confrontations that need to be resolved (Carvalho et al. 2010). One of the most profitable ways



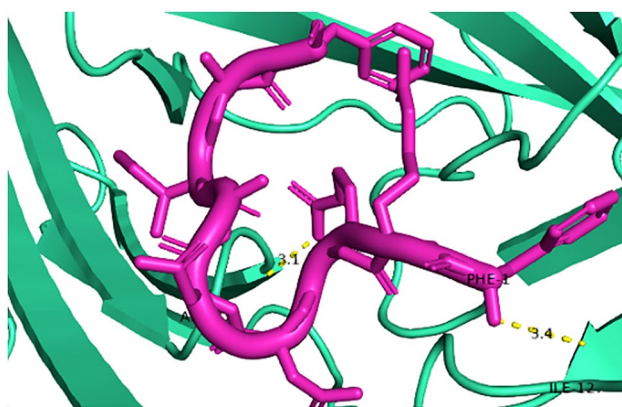
**Fig. 2** This docking study shows the interaction analysis of epitope YHNMSPNIR (represented with light magenta color) with 4AH2 receptor (described with green cyan color). It shows that the epitope gets hold of the 4AH2 receptor with the help of 3 Hydrogen bonds (Color figure online)



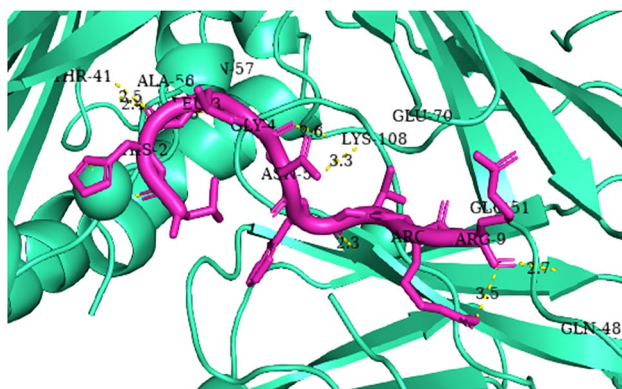
**Fig. 3** This docking study shows the interaction analysis of epitope LNLSPLIQR (represented with light magenta color) with 6CQL receptor (described with green cyan color). It shows that the epitope gets to the 6CQL receptor with the help of 5 Hydrogen bonds (Color figure online)

to confront the use of immunoinformatics (Kardani et al. 2020). Immunoinformatics can play a significant role in developing vaccines regarding diseases that are threatening humanity and society. It mainly focuses on using bioinformatics tools and software to predict the antigenic sequences that can provoke a better immune response in the host. B cell epitopes are critical in developing in silico vaccines and are believed to elicit a proper adaptive immune response. It





**Fig. 4** This docking study shows the interaction analysis of epitope FRRNNAIAF (represented with light magenta color) with 3C5J receptor (described with green cyan color). It shows that the epitope gets hold of the 3C5J receptor with the help of 2 Hydrogen bonds (Color figure online)



**Fig. 5** This docking study shows the interaction analysis of epitope LHLGNFVRR (represented with light magenta color) with 6CQL receptor (described with green cyan color). It shows that the epitope gets hold of the 6CQL receptor with the help of 10 Hydrogen bonds (Color figure online)

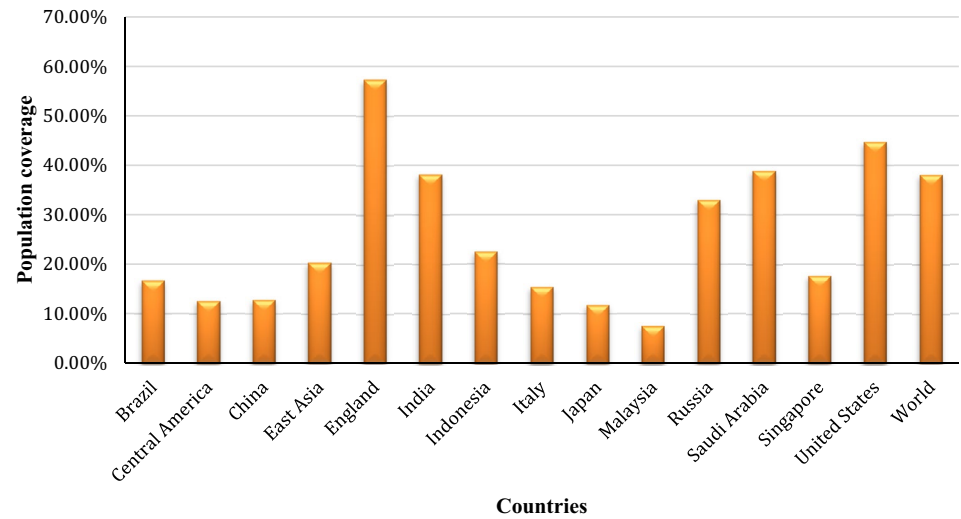
**Table 5** List showing number of HLA binders and analysis of their Binding affinity as analyzed by MHCpred tool

Epitope	Number of HLA binders	HLA with predicted IC <sub>50</sub> nM value
YHNMSPNIR	1	HLA-DRB1_0101 (49.31)
LNLSPLIQR	2	HLA-DRB1_0101 (30.41) HLA-DRB1_0401 (413.05)
FRRNNAIAF	3	HLA-DRB1_0101 (5.24) HLA-DRB1_0401 (424.62) HLA-DRB1_0701 (21.83)
LHLGNFVRR	2	HLA-DRB1_0101 (299.23) HLA-DRB1_0701 (359.75)

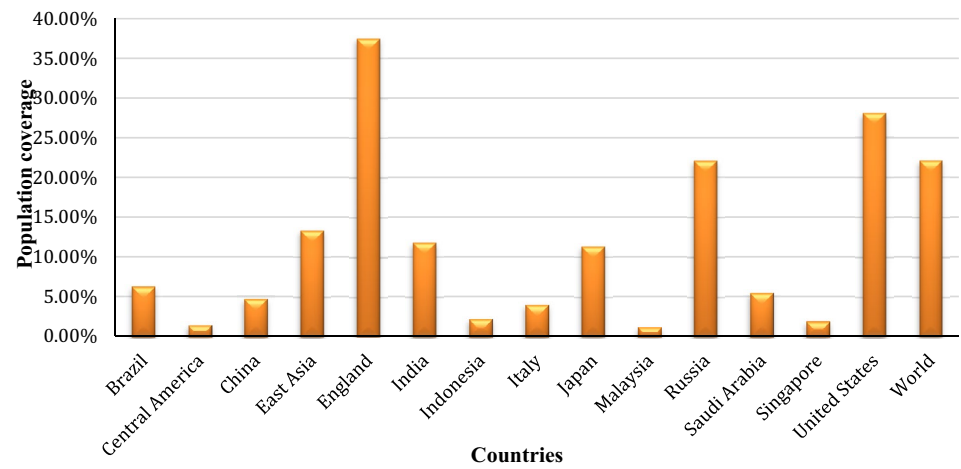
Good binders show IC<sub>50</sub> value < 500 nM

is believed that comprehensive computational screening of the proteome of the NV-B strain of the Nipah virus had led us to the most capable and potential epitopes that can play a better role among others to elicit a robust immune response against the causative agent. After molecular docking, the selection of four epitopes (LHLGNFVRR, LNLSPLIQR, YHNMSPNIR, and FRRNNAIAF) was accomplished using a low binding affinity factor. Analysis of population coverage worldwide was also implemented on the anticipated epitopes. It was found that epitope LHLGNFVRR can be predicted as the most promising vaccine candidate because of its low binding energy and high population coverage globally. The final selected peptide was nontoxic as well as non-allergic. It could bind to HLA alleles with strong affinity. Finally, simulation validates the efficiency of peptide—allele stability complex. Wet-lab experiments followed by clinical trials can further verify this. These *insilico* ways of designing a vaccine are powerful tools to combat emerging diseases like the Nipah virus.

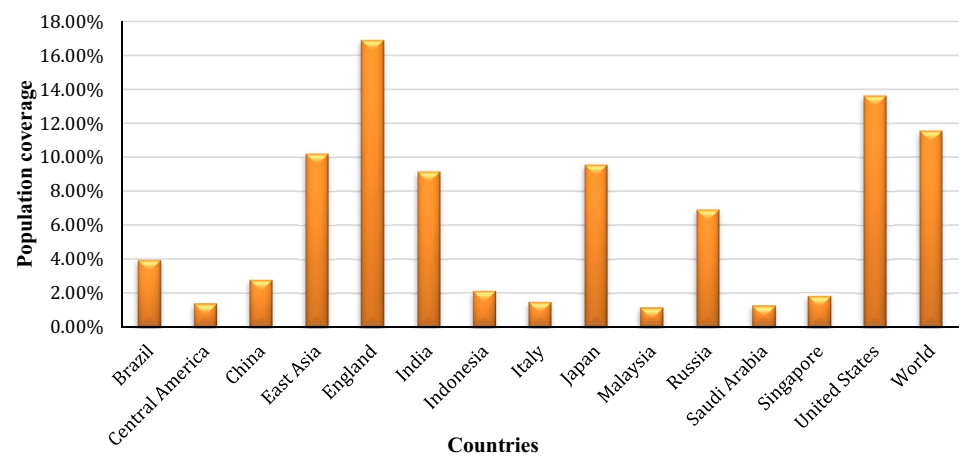
**Fig. 6** Graphical representation of population coverage analysis of epitope FRRNNAIAF across different regions of the world



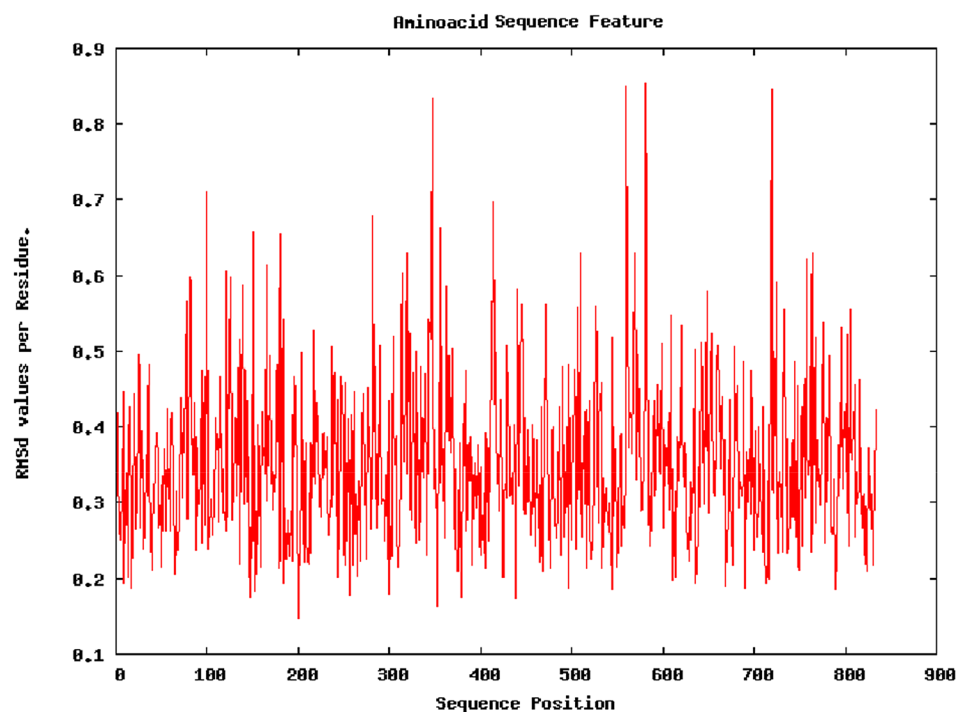
**Fig. 7** Graphical representation of population coverage analysis of epitopes LNLSPLIQR and LHLGNFVRR across different world regions



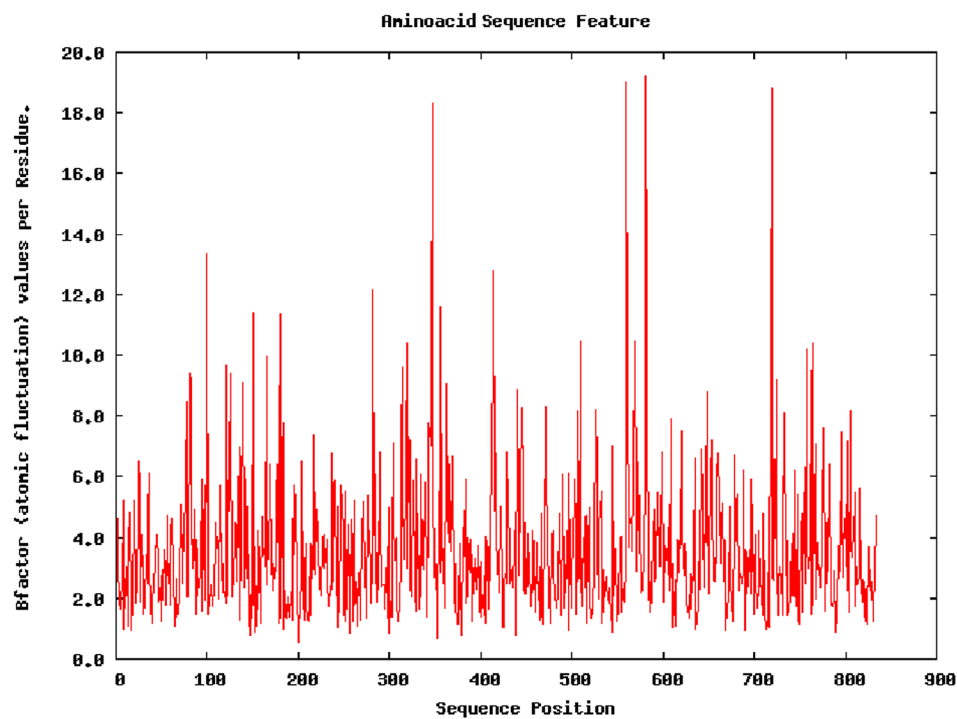
**Fig. 8** Graphical representation of population coverage analysis of epitope YHNMPSNIR across different regions of the world



**Fig. 9** Graph displaying RMSD for epitope LHLGNFVRR with 6CQL receptor obtained during performing simulation studies



**Fig. 10** B factor plot of epitope LHLGNFVRR with 6CQL receptor obtained during performing simulation studies



## Declarations

**Conflict of interest** The authors, as a result of this, declare they have no conflict of interest.

**Ethical Approval** The authors did not perform any experiments on humans or animals.

## References

- Aadit, Shariff M (2019) Nipah virus infection: a review. *Epidemiol Infect* 147:e95
- Aljofan M (2013) Hendra and Nipah infection: emerging paramyxoviruses. *Virus Res* 177(2):119–126
- 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 Bioinform* 7:153
- Carvalho JA, Rodgers J, Atouguia J, Prazeres DM, Monteiro GA (2010) DNA vaccines: a rational design against parasitic diseases. *Expert Rev Vaccines* 9(2):175–191
- CDC (2017). <https://www.cdc.gov/vhf/Nipah/index.html>. Accessed 21 Apr 2020
- CDC (2019). <https://www.cdc.gov/globalhealth/healthprotection/fieldupdates/fall-2018/india-Nipah-response.html>. Accessed 13 July 2020
- Chadha MS, Comer JA, Lowe L, Rota PA, Rollin PE, Bellini WJ, Ksiazek TG, Mishra AC (2006) Nipah virus-associated encephalitis outbreak Siliguri India. *Emerg Infect Dis*. <https://doi.org/10.3201/eid1202.051247>
- Chong HT, Kamarulzaman A, Tan CT, Goh KJ, Thayaparan T, Kunjapan SR et al (2001) Treatment of acute Nipah encephalitis with ribavirin. *Ann Neurol* 49(6):810–813
- Clayton BA, Middleton D, Bergfeld J, Haining J, Arkinstall R, Wang L, Marsh GA (2012) Transmission router for Nipah virus from Malaysia and Bangladesh. *Emerg Infect Dis* 18:1983–1993
- De Groot AS, Jesdale B, Martin W, Saint AC, Sbai H, Bosma A et al (2003) Mapping cross-clade HIV-1 vaccine epitopes using a bioinformatics approach. *Vaccine* 21(27–30):4486–4504
- De Wit E, Munster VJ (2014) Animal models of disease shed light on Nipah virus pathogenesis and transmission. *J Pathol* 235(2):196–205
- Doytchinova IA, Flower DR (2007) VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit Vaccines. *BMC Bioinform* 8:4
- Epstein JH, Field HE, Luby S, Pulliam JR, Daszak P (2006) Nipah virus: impact, origins, and causes of emergence. *Curr Infect Dis Rep* 8:59–65
- Goh KJ, Tan CT, Chew NK, Tan PS, Kamarulzaman A, Sarji SA, Wong KT, Abdullah BJ, Chua KB, Lam SK (2000) Clinical features of Nipah virus encephalitis among pig farmers in Malaysia. *N Engl J Med* 342:1229–1235
- Griffin BD, Leung A, Chan M, Warner BM, Ranadheera C, Tierney K, Audet J, Frost KL, Safronetz D, Embury-Hyatt C, Booth SA, Kobasa D (2019) Establishment of an RNA polymerase II-driven reverse genetic system for Nipah virus strain from Malaysia and Bangladesh. *Sci Rep* 10(1):38–43
- Guan P, Doytchinova IA, Zygouri C, Flower DR (2003) MHCpred: a server for quantitative prediction of peptide–MHC Binding. *Nucleic Acid Res* 31:3621–3624
- Gupta S, Kapoor P, Chaudhary K, Gautam A, Kumar R, Raghava GPS (2013) In silico approach for predicting toxicity of peptides and proteins. *PLoS ONE* 8(9):e73957
- Harcourt BH, Lowe L, Tamin A, Liu A, Bankamp B, Bowden N, Rollin PE, Comer JA, Ksiazek TG, Gurley ES, Breiman RF, Bellini WJ, Rota PA, Hossain MJ (2004) Genetic characterization of Nipah virus, Bangladesh. *Emerg Infect Dis* 11(10):1594–1597
- Hospital A, Andrio P, Fenollosa C, Cicin-Sain D, Orozco M, Gelpí JL (2012) MDWeb and MDMoby: an integrated web based platform for molecular dynamics simulation. *Bioinformatics* 28:1278–1279
- Hossain MJ, Gurley ES, Montgomery JM, Bell M, Carroll DS, Hsu VP, Formenty P, Croisier A, Bertherat E, Faiz MA et al (2008) Clinical representation of nipah virus infection in Bangladesh. *Clin Infect Dis* 46:977–984
- Hsu VP, Hossain MJ, Parashar UD, Ali MM, Ksiazek TG, Kuzmin I, Niezgoda M, Rupprecht C, Bresee J, Breiman RF (2005) Nipah virus encephalitis reemergence, Bangladesh. *Emerg Infect Dis*. <https://doi.org/10.3201/eid1012.040701>
- Jensen KK, Andreatta M, Marcatili P, Buus S, Greenbaum JA, Yan Z, Sette A, Peters B, Nielsen M (2018) Improved methods for predicting peptide binding affinity to MHC class II molecules. *Immunology* 8:4
- Kardani K, Bolhassani A, Namvar A (2020) An overview of in silico vaccine design against different pathogens and cancer. *Expert Rev Vaccines* 19(8):699–726
- Kazi A, Chuah C, Majeed BA, Leow CH, Lim BH, Leow CY (2018) Current progress of immunoinformatics approach harnessed for cellular and antibody dependent vaccine design. *Pathog Glob Health*. <https://doi.org/10.1080/20477724.2018.1446773>
- King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ (2011) Virus taxonomy: classification and nomenclature of viruses—Ninth Report of the International Committee on Taxonomy of Viruses. Academic Press, Amsterdam
- Kulkarni DD, Tosh C, Venkatesh G, Senthil DK (2013) Nipah virus infection: current scenario. *Indian J Virol* 24(3):398–408
- Liu Q, Bradel-Tretheway B, Monreal AI et al (2015) Nipah virus attachment glycoprotein stalk C-terminal region links receptor binding to fusion triggering. *J Virol* 89(3):1838–1850
- Lo MK, Miller D, Aljofan M, Mungall BA, Rollin PE, Bellini WJ et al (2010) Characterization of the antiviral and inflammatory responses against Nipah virus in endothelial cells and neurons. *Virology* 404:78–88
- Lo MK, Feldmann F, Gary JM, Jordan R, Bannister R, Cronin J, de Wit E (2019) Remdesivir (GS-5734) protects African green monkeys from Nipah virus challenge. *Sci Transl Med* 11(494):eaau242
- Manzoni C, Kia DA, Vandrovcova J, Hardy J, Wood NW, Lewis PA, Ferrari R (2018) Genome transcriptome and proteome the rise of omics data and their integration in biomedical sciences. *Brief Bioinform* 19(2):286–302
- Mazzola LT, Kelly-Cirino C (2019) Diagnostics for Nipah virus: a zoonotic pathogen endemic to Southeast Asia. *BMJ Glob Health* 4:e001118
- Meza B, Ascencio F, Sierra-Beltrán AP, Torres J, Angulo CA (2017) Novel design of a multi-antigenic, multistage and multi-epitope vaccine against *Helicobacter pylori*: an in silico approach. *Infect Genet Evol* 49:309–317
- Nezafat N, Karimi Z, Eslami M, Mohkam M, Zandian S, Ghasemi Y (2016) Designing an efficient multi-epitope peptide vaccine against *Vibrio cholerae* via combined immunoinformatics and protein interaction based approaches. *Comput Biol Chem* 62:82–95
- Peng M, Mo Y, Wang Y et al (2019) Neoantigen vaccine: an emerging tumor immunotherapy. *Mol Cancer* 18:129
- Rahman MA, Hossain MJ, Sultana S, Homaira N, Khan SU, Rahman M, Gurley ES, Rollin PE, Lo MK, Comer JA et al (2012) Date palm sap linked to Nipah virus outbreak in Bangladesh, 2008. *Vector Borne Zoonotic Dis* 12:65–72
- Saha S, Raghava GPS (2006) AllgPred: prediction of allergenic proteins and mapping of IgE epitope. *Nucleic Acid Res* 34:w202–209



- Sazzad HM, Hossain MJ, Gurley ES, Ameen KM, Parveen S, Islam MS, Faruque LI, Podder G, Banu SS, Lo MK et al (2013) Nipah virus infection outbreak with nosocomial and corpse-to-human transmission, Bangladesh. *Emerg Infect Dis* 19:210–217
- Schiavina M, Salladini E, Murrall MG, Tria G, Felli IC, Pierattelli R, Longhi S (2020) Ensemble description of the intrinsically disordered N-terminal domain of the Nipah virus P/V protein from combined NMR and SAXS. *Sci Rep* 10:19574
- Singh S, Singh H, Tuknait A, Chaudhary K, Singh B, Kumaran S, Raghava GPS (2015) PEPstrMOD: structure prediction of peptides containing natural, non-natural and modified residues. *BMC Biol Direct* 10:73
- Trott O, Olson AJ (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimisation and multithreading. *J Comput Chem*. <https://doi.org/10.1002/jcc.21334>
- Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer FT, de Beer TAP, Rempfer C, Bordoli L, Lepore R, Schwede T (2018) SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acid Res* 46:w296–w303
- WHO (2018). <https://www.who.int/csr/don/07-august-2018-Nipah-virus-india/en/>. Accessed 20 June 2020
- WHO (2019). <https://www.who.int/southeastasia/outbreaks-and-emergencies>. Accessed 15 May 2020
- Wong KT, Shieh WJ, Kumar S et al (2002) Nipah virus infection: pathology and pathogenesis of an emerging paramyxoviral zoonosis. *Am J Pathol* 161:2153–2167
- Zhang L (2017) Multi-epitope vaccines: a promising strategy against tumors and viral infections. *Cell Mol Immunol* 15(2):182–184
- Zhou WY, Shi Y, Wu C, Zhang WJ, Mao XH, Guo G, Zou QM (2009) Therapeutic efficacy of a multi-epitope vaccine against *Helicobacter pylori* infection in BALB/c mice model. *Vaccine* 27(36):5013–5019

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