HOSTED BY
ELSEVIER

Contents lists available at ScienceDirect

Asian Pacific Journal of Tropical Medicine

journal homepage: http://ees.elsevier.com/apjtm

Original research http://dx.doi.org/10.1016/j.apjtm.2017.06.016

In silico identification and characterization of common epitope-based peptide vaccine for Nipah and Hendra viruses

Chayan Kumar Saha[1], Md. Mahbub Hasan[2], Md. Saddam Hossain[1], Md. Asraful Jahan[1], Abul Kalam Azad[1]✉

[1]*Department of Genetic Engineering & Biotechnology, Shahjalal University of Science and Technology, Sylhet 3114, Bangladesh*

[2]*Department of Genetic Engineering & Biotechnology, University of Chittagong, Chittagong 4331, Bangladesh*

ARTICLE INFO

Article history:
Received 19 Nov 2016
Received in revised form 10 Apr 2017
Accepted 28 May 2017
Available online 1 Jul 2017

Keywords:
Hendra virus
Nipah virus
Epitope
Encephalitis
Docking
Flanking dipeptides

ABSTRACT

Objective: To explore a common B- and T-cell epitope-based vaccine that can elicit an immune response against encephalitis causing genus *Henipaviruses*, Hendra virus (HeV) and Nipah virus (NiV).

Methods: Membrane proteins F, G and M of HeV and NiV were retrieved from the protein database and subjected to different bioinformatics tools to predict antigenic B-cell epitopes. Best B-cell epitopes were then analyzed to predict their T-cell antigenic potentiality. Antigenic B- and T-cell epitopes that shared maximum identity with HeV and NiV were selected. Stability of the selected epitopes was predicted. Finally, the selected epitopes were subjected to molecular docking simulation with HLA-DR to confirm their antigenic potentiality *in silico*.

Results: One epitope from G proteins, one from M proteins and none from F proteins were selected based on their antigenic potentiality. The epitope from the G proteins was stable whereas that from M was unstable. The M-epitope was made stable by adding flanking dipeptides. The 15-mer G-epitope (VDPLRVQWRNNSVIS) showed at least 66% identity with all NiV and HeV G protein sequences, while the 15-mer M-epitope (GKLEFRRNNAIAFKG) with the dipeptide flanking residues showed 73% identity with all NiV and HeV M protein sequences available in the database. Molecular docking simulation with most frequent MHC class-II (MHC II) and class-I (MHC I) molecules showed that these epitopes could bind within HLA binding grooves to elicit an immune response.

Conclusions: Data in our present study revealed the notion that the epitopes from G and M proteins might be the target for peptide-based subunit vaccine design against HeV and NiV. However, the biochemical analysis is necessary to experimentally validate the interaction of epitopes individually with the MHC molecules through elucidation of immunity induction.

1. Introduction

A consequence of Nipah virus (NiV) infection in humans is encephalitis which causes headache, stiff neck, fever and altered mental status due to inflammation of parts of the brain, spinal cord and meninges [1]. Since 1994, there have been 31 outbreaks of Hendra virus (HeV); 17 of these outbreaks occurred in 2011. In 1998–1999, the first outbreak of NiV was documented in Malaysia and Singapore, where 283 cases of NiV encephalitis were observed, with 109 fatalities [2]. To date, there have been a total of 600 reported cases of NiV infection in people with fatality rates up to 100% for some outbreaks [3]. Most importantly, in Bangladesh, outbreaks occurred almost every year between 2001 and 2015, which have had higher fatality rate averaging about 75% [4]. In January, 2012, NiV outbreak occurred in the Joypurhat district of Bangladesh, resulting in the death of six individuals, with case fatality rate 100% indicating the continuous role of this virus in disease outbreaks in Bangladesh [5]. In the past, NiV showed

First author: Chayan Kumar Saha, Department of Genetic Engineering & Biotechnology, Shahjalal University of Science and Technology, Sylhet 3114, Bangladesh.
Tel: +88 01717650469
E-mail: chayan.sust7@gmail.com
✉Corresponding author: Abul Kalam Azad, Department of Genetic Engineering & Biotechnology, Shahjalal University of Science and Technology, Sylhet 3114, Bangladesh.
Tel: +88 0821 717850x411
Fax: +88 0821 725050
E-mail: dakazad-btc@sust.edu
Peer review under responsibility of Hainan Medical University.

transmission from bats to pigs and then to humans suggesting that the pig was required as amplifying host. Recent outbreaks in Bangladesh from 2001 to 2005 reported transmission of NiV directly from bats to humans and even human-to-human [6]. NiV belongs to the order *Mononegavirales* and is a member of the *Paramyxoviridae* family. Because of unique biological and genetic characteristics, NiV and the closely related HeV form a new genus *Henipavirus* within the subfamily *Paramyxoviridae* [7]. HeV, like NiV, is a zoonotic pathogen containing non-segmented, negative-stranded RNA genome [5]. To date, four of the seven human infections with HeV recorded so far have been fatal [5]. In nature, both NiV and HeV are found in several pteropous fruit bat species (flying foxes) and possess a broad host tropism that includes pigs, horses, cats, dogs, guinea pigs, hamsters, ferrets, monkeys and humans. Unlike NiV, HeV is transmitted from bats-to-horses and then to humans [3]. Both viruses can cause relapsed encephalitis in infected individuals. The case fatality rate of human infection varies between 50% and 100%, made both NiV and HeV, the most deadly viruses known to infect humans [5].

There is currently no vaccine or antiviral therapy licensed for use in humans or animals. During the initial NiV outbreak in Malaysia, ribavirin used against NiV, did not show much success [8]. In Australia, ribavirin failed to show a good therapeutic effect in several patients infected with HeV [9]. Additionally, treatment with chloroquine and ribavirin proved ineffective for one HeV-infected individual in 2009 as no clinical benefit was observed [10]. Therefore, a vaccine against both NiV and HeV should be designed that will be safe and effective for using in humans. *In silico* design of epitope based vaccines has become the standard approach to vaccine discovery in the current post-genomic era as these offers inexpensive, more specific, potential and easy means of vaccine development for treating infectious diseases [11]. Suitable antigenic determinants of fusion protein (F), attachment glycoprotein (G) and matrix protein (M) might be potential target to design epitope-based vaccine against encephalitis caused by NiV and HeV infection. In the present study, we analyzed F, G and M proteins of HeV and NiV by using different bioinformatics tools to identify the best antigen that might be an effective vaccine target against both HeV and NiV. We conducted investigations to find the best B- and T-cell epitopes from F, G and M proteins of NiV and HeV following antigenicity analysis and showed the binding ability of the epitopes with MHC class-I (MHC I) and class-II (MHC II) by *in silico* molecular docking approach.

2. Methodology

2.1. Prediction of antigenicity of different Henipavirus membrane proteins

All the complete amino acid sequences of each F, G and M proteins of NiV and HeV were retrieved from protein databases (http://www.us.expasy.org/sprot; http://www.ncbi.nlm.nih.gov/protein/) and non-identical sequences were analyzed with VaxiJen v2.0 (http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html) antigen prediction server [12]. For highest accuracy, a threshold value of 0.5 was used to check the antigenicity of each full length protein. Amino acid sequences from F and G Proteins that have antigenic score >0.5 were chosen. Only for the M proteins that have antigenic score >0.4 were selected as described previously [13]. Each chosen full length amino acid sequence was then subjected to transmembrane topology analysis using TMHMM v.2.0 (http://www.cbs.dtu.dk/services/TMHMM/) prediction server to identify exo-membrane (surface exposed) amino acid sequences of each protein [14].

2.2. Prediction of antigenic B-cell epitopes

For the prediction of B-cell epitopes, each selected full length protein sequence was subjected to BCPreds (http://ailab.cs.iastate.edu/bcpreds/predict.html) analysis where both BCpred and AAP prediction methods were used [15]. All predicted B-cell epitopes (16-mer) having a BCPreds cutoff score >0.8 were selected and subsequently checked for membrane topology by comparing with TMHMM results for exo-membrane amino acid sequences. Surface exposed B-cell epitope sequences having the cutoff value >0.8 for BCPreds were then analyzed using VaxiJen at threshold 0.5 to check the antigenicity. Finally, two-three epitopes with the top VaxiJen scores were selected to use in prediction of T-cell epitopes.

2.3. Prediction of T-cell epitopes from selected B-cell epitopes

T-cell epitopes were predicted from the selected B-cell epitopes and two screening steps were followed. In the first screening, the selection criteria were: i) the T-cell epitope sequence should bind to both the MHC I and MHC II molecules and the least number of total interacting MHC molecules should be >15, ii) the T-cell epitope sequence must interact with HLA-DRB1*0101 of MHC-II, and iii) T-cell epitope sequence should be antigenic based on VaxiJen score. Propred-1 (http://www.imtech.res.in/raghava/propred1/) [16] and Propred (http://www.imtech.res.in/raghava/propred/) [17] servers that utilize amino acid position coefficients inferred from literature employing linear prediction model [18], were used to identify common epitopes that bind to both MHC I and MHC II molecules. Total numbers of interacting MHC alleles were counted. For quantitative structure–activity relationship (QSAR) simulation approach, the half maximal (50%) inhibitory concentration (IC_{50}) and antigenicity of common epitopes (predicted by Propred-1 and Propred) were calculated using MHCPred v.2 (http://www.ddg-pharmfac.net/mhcpred/MHCPred/) server [19] and VaxiJen v2.0, respectively. Epitopes with highest antigenicity and those bound more than 15 molecules of both MHC I and MHC II alleles and had less than 100 nmol/L IC_{50} scores for DRB1*0101 were selected. The second screening was based on structure and QSAR simulation methods using T-Epitope Designer (http://www.bioinformation.net/ted/) [20] and MHCPred, respectively. T-epitope Designer can screen peptides for >1000 HLA alleles. In the second screening, the criteria were as follows. The peptide should bind >75% of total HLA molecules; the peptide must bind with high scores to (i) HLA-A*0201, HLA-A*0204, and HLA-B*2705, and (ii) DRB1*0101 and DRB1*0401. T-epitope Designer was used for first two criteria and MHCPred was used for the final criteria. The final list of T-cell epitopes was made with peptide sequences that pass these above mentioned criteria and VaxiJen

and IC_{50} scores. Physico-chemical properties of identified B- and T-cell epitopes (15-mer) were analyzed with ProtParam computer program (http://web.expasy.org/protparam/).

2.4. Molecular docking of predicted epitopes with HLA-DR

The 3D structures of B- and T-cell epitopes was built by PyMol molecule builder [21] and subjected for molecular docking simulation with MHC I molecules (HLA-B7, HLA-B27, HLA-B44, and HLA-B62) and MHC II molecules (HLA-DRB1*0101, HLA-DRB1*0301, HLA-DRB1*0401, HLA-DRB1*1501, HLA-DRB1*0701, HLA-DRB1*1101 and HLA-DRB1*1801 and HLA-DRA) by Molsoft ICM-pro 3.5 [22]. Similarity among MHC I molecules (HLA super types B7, B27, B44 and B62) were analyzed using multiple sequence alignment (http://www.ebi.ac.uk/Tools/msa/clustalw2/). In Protein Data Bank (PDB) (http://www.rcsb.org/pdb/home/home.do), there was no experimental structure available for HLA-B*2702 and HLA-B*5201 which are representative of MHC I HLA super types B27 and B62, respectively. The 3D structures of these two proteins were obtained using homology modeling with SWISS-MODEL (http://swissmodel.expasy.org) [23]. The homology models of HLA-B*2702 (Uniprot KB ID P03989) and HLA-B*5201 (Uniprot KB ID P30490) were constructed based on the experimental structure of HLA-B*2709 A chain (PDB ID 1K5N-a) and HLA-B*5101 A chain (PDB ID 1E27-a), respectively. The amino acid sequence identity of HLA-B*2702 and HLA-B*5201 with the A chain of both HLA-B*2709 and HLA-B*5101 were 98.551% and 99.275%, correspondingly. The stereo chemical quality of the homology models were assessed by the PROCHECK program through PDBsum server (http://www.ebi.ac.uk/pdbsum/) [24] (Table 1). PDB ID for experimental structures of HLA-B*5301, representative of HLA super type B7 is 1A1M-a and that of HLA-B*4403, representative of HLA super type B44 is 3KPN-a. The PDB ID of HLA-DRB1*0101, HLA-DRB1*0301, HLA-DRB1*0401, HLA-DRB1*1501 and HLA-DRA were 2FSE-b, 1A6A-b, 1J8H-b, 1BX2-b and 1A6A-a, respectively. According to the Entrez cross-database search page (http://www.ncbi.nlm.nih.gov/sites/gquery), HLA-DRB1*0701, HLA-DRB1*1101 and HLA-DRB1*1801 have the same 3D structure (PDB ID, 3QXA-b) and for this, molecular docking of the epitope was conducted against five HLA-DRB 3D structures. Similarity among those five HLA-DRB molecules was analyzed using multiple sequence alignment (http://www.ebi.ac.uk/Tools/msa/clustalw2/). The best docking orientation was selected based on binding free energy and hydrogen bond distance. The *in silico* approaches for identification and characterization of the epitopes were validated as described previously [25].

Table 1
Ramachandran plot statistics of homology models of HLA-B2702 and HLA-B5201 computed with PROCHECK program.

Plot statistics	HLA-B2702	HLA-B5201
% residues in favorable regions	90.9	90.1
% residues in additional residue regions	8.3	8.7
% residues in generously regions	0.4	0.8
% residues in disallowed regions	0.4	0.4
% of non-proline and non-glycine residues	100.0	100.0

3. Results

3.1. Antigenicity and topology of selected proteins

NiV and HeV have 68%–92% amino acid homology in the protein-coded regions and 40%–67% nucleotide homology in the un-translated regions [26]. The homology of F, G and M proteins is approximately 88%, 83% and 89% between NiV and HeV [27]. Due to high percentage of identity at amino acid level between these two viruses, the present study focuses on identification of an efficient common epitope for NiV and HeV. The antigenicity was identified based on auto cross covariance (ACC) transformation of protein sequences using VaxiJen. An analysis of exo-membrane full length proteins selected for this study exhibited various degree of antigenicity. The highest VaxiJen score indicating the highest antigenicity of an amino acid sequence of F, G and M proteins of NiV was 0.502, 0.525 and 0.410, respectively and that of an amino acid sequence of those proteins of HeV was 0.552, 0.551 and 0.442, correspondingly (Table 2). Although NiV M and HeV M proteins exhibited the VaxiJen score less than 0.5 and predicted to

Table 2
Accession numbers, VaxiJen score and exo-membrane sequences of F, G and M proteins of NiV and HeV.

Protein	Accession number	VaxiJen score	Exo-membrane sequence
NiV F protein (NiV F)	ACT32614	0.488	131–495
	AAK50553	0.501	131–495
	AEZ01396	0.489	131–495
	CBM41033	0.495	131–495
	AAY43915	0.486	131–495
	AAM13405	0.502	131–495[a]
	CAF25496	0.500	131–495
HeV F protein (HeV F)	AAB39505	0.546	30–495
	NP_047111	0.552	30–495[a]
	AEB21197	0.552	30–495
	AEQ38114	0.549	30–495
NiV G protein (NiV G)	AAK29088	0.510	70–602
	CAF25497	0.518	70–602
	ADN51995	0.525	70–602
	AAX51853	0.517	70–602
	AAY43916	0.523	70–602
	ACT32615	0.525	70–602[a]
	AEZ01397	0.520	70–602
HeV G protein (HeV G)	AEB21225	0.551	70–604[a]
	NP_047112	0.538	70–604
	AEQ38115	0.536	70–604
	AEB21206	0.538	70–604
	AEQ38026	0.536	70–604
	AEB21216	0.524	70–604
	AEQ38108	0.538	70–604
NiV M protein (NiV M)	Q9IK90	0.401	1–352
	D2DEB9	0.409	1–352
	Q4VCP7	0.409	1–352
	E4Z9G1	0.410	1–352[a]
	CAF25495	0.407	1–352
HeV M protein (HeV M)	AAB39504	0.414	1–352
	AEB21223	0.428	1–352
	NP_047110	0.442	1–352[a]
	AEB21208	0.439	1–352
	AEQ38060	0.427	1–352
	AEB21196	0.429	1–352

Proteins with highest VaxiJen score indicating highest antigenicity are shown with superscript letter a.

Table 3
B-cell epitopes from full length F, G and M proteins of NiV and HeV.

Protein name	Serial no. of B-cell epitopes	Amino acid start position of epitope	BCPred epitope sequences	BCPred scores	VaxiJen Scores	Exo-membrane sequences
NiV F	1	293	IQELLPVSFNNDNSEW	1.000	0.827	131–495
	2	298	PVSFNNDNSEWISIVP	0.965	0.988	131–495
	3	435	SVNYNSEGIAIGPPVF	1.000	0.795	131–495
HeV F	1	293	VQELLPVSFNNDNSEW	1.000	0.827	30–495
	2	298	PVSFNNDNSEWISIVP	0.965	0.988	30–495
	3	408	TLLMIDNTTCTTVVLG	0.998	0.693	30–495
NiV G	1	107	PKVSLIDTSSTITIPA	0.933	0.683	70–602
HeV G	1	162	CPNPLPFREYRPISQG	1.000	1.319	70–604
	2	205	YTLPINTREGVCITDP	1.000	1.078	70–604
	3	472	VDPLRVQWRNNSVISR	1.000	1.233	70–604
NiV M	1	189	IPRTMLEFRRNNAIAF	1.000	1.002	1–352
	2	190	PRTMLEFRRNNAIAFN	0.956	1.073	1–352
NiV M	1	189	IPRTMLEFRRNNAIAF	1.000	1.002	1–352
	2	190	PRTMLEFRRNNAIAFN	0.956	1.073	1–352

B-cell epitopes were predicted in F, G and M proteins of NiV and HeV having the highest VaxiJen score using BCPred and AAP. Antigenicity of B-cell epitopes derived from each protein was calculated using VaxiJen. Exo-membrane amino acid sequence of the proteins is mentioned.

Table 4
T-cell epitopes predicted from selected B-cell epitopes mentioned in Table 3.

Protein	Serial no. of B-cell epitope	Predicted T-cell epitopes	VaxiJen score	Amino acid start position	IC_{50} Value of T-cell epitopes for DRB1*0101 (MHCPred)	Number of MHC binding alleles		
						MHC I (Propred-I)	MHC II (Propred)	Total
HeV G	3	LRVQWRNNS	1.840	475	7.85	2	33	35
NiV M	1	FRRNNAIAF	1.786	196	5.24	6	28	34
	2	LEFRRNNAI	0.775	194	88.31	13	3	16
HeV M	1	FRRNNAIAF	1.786	196	5.24	6	28	34
	2	LEFRRNNAI	0.775	194	88.31	13	3	16

Common T-cell epitopes predicted with MHCPred, Propred-I and Propred are shown with their various parameters.

be non-antigenic by VaxiJen, they were also considered for further analysis as described previously [13]. The fundamental criterion of a good epitope is that it must be exposed to cell outside. The transmembrane topology analysis of these proteins were done using TMHMM and the result revealed that the lengths of exo-membrane amino acid sequences for F, G and M proteins of NiV were 131–495, 70–602 and 1–352, respectively and that of those proteins of HeV were 30–495, 70–604 and 1–352, correspondingly (Table 2).

3.2. Antigenic B-cell epitopes

For becoming a good vaccine candidate, a peptide should be hydrophilic and produce both the B-cell and T-cell mediated immunity [16]. Therefore, full length proteins were first analyzed for B-cell epitope prediction using BCpreds and all predicted B-cell epitopes were listed from each protein (Table 3). Best epitopes were selected based on the criteria as mentioned in methods. In general, 16-mer epitopes having BCpreds [15] and VaxiJen cutoff values >0.8 and >0.5, respectively were selected. Considering the BCpreds and VaxiJen cutoff values, total 14 epitopes (three 16-mer epitopes from each of NiV F, HeV F and HeV G and two 16-mer epitopes from each of NiV M and HeV M, one from NiV G) were finally selected for further analysis as B-cell epitope (Table 3). Each selected B-cell epitope from a protein, showed 100% identity with all retrieved amino acid sequences of that protein.

3.3. B-cell epitopes derived T-cell epitopes

Each B-cell epitope selected in Table 3 was analyzed to identify T-cell epitope. In first level screening, analysis with Propred-I, Propred and MHCPred identified three 9-mer epitopes (LRVQWRNNS, FRRNNAIAF and LEFRRNNAI), which were all antigenic with good VaxiJen score (Table 4). The result showed that the total number of interacting MHC molecules were more than 15 and interaction with HLA-DRB1*0101 of MHC II having less IC_{50} score than 100 nmol/L was predicted (Table 4). Propred and Propred-1, two server based tools were used for *in silico* T-cell epitope prediction [28]. At the second

Table 5
Final selection of T-cell epitopes.

Protein	Epitope	T-epitope designer score of T-cell epitopes			IC_{50} Value of T-cell epitopes	
		A*0201	A*0204	B*2705	DRB1*0101 (MHCPred)	DRB1*0401 (MHCPred)
HeV G	LRVQWRNNS	1137.01	643.59	1930.42	7.85	1196.74
HeV/NiV M	FRRNNAIAF	214.14	303.65	183.04	5.24	424.62
HeV/NiV M	LEFRRNNAI	877.46	686.96	1351.80	88.31	1999.86

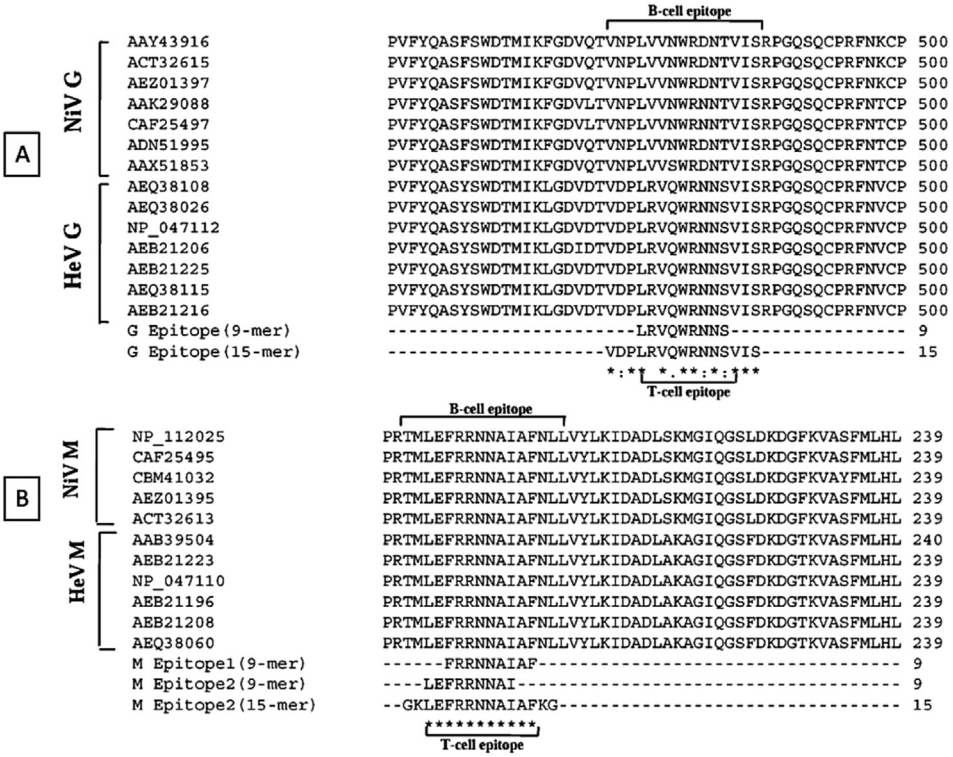

Figure 1. Multiple alignment of MHC I (HLA super types B7, B27, B44, B62) (A) and MHC II (B) molecules.

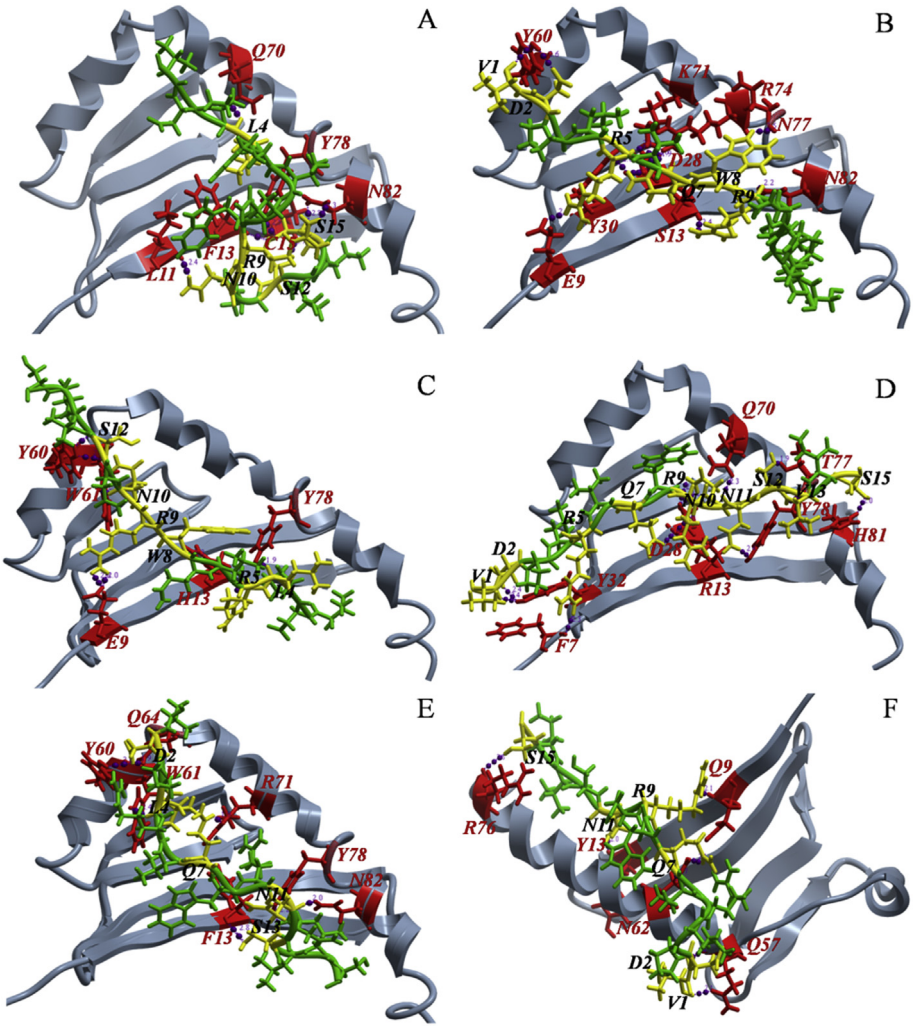

Figure 2. Molecular docking simulation of the G-epitope with MHC II molecules HLA-DRB1*0101 (A), HLA-DRB1*0301 (B), HLA-DRB1*0401 (C), HLA-DRB1*1501 (D), HLA-DRB1*0701 (E) and HLA-DRA (F).

MHC II structures are shown as gray ribbon and the amino acid residues involved in the H-bonding are shown as red sticks and labeled red. The epitope is shown as green ribbon and the amino acid residues involved in the H-bonding are shown as yellow sticks and labeled black. H-bonds are displayed as blue spheres and H-bonding distances are labeled blue.

level of screening, identified three peptides (LRVQWRNNS, FRRNNAIAF and LEFRRNNAI) from the first screening were used to predict their binding abilities to >1000 MHC alleles using T-Epitope Designer and epitopes that bound to >75% MHC alleles were selected. Similarly, for mostly used A*0201, A*0204, and B*2705 alleles, the cutoff value was set in a way that selected peptides must bind to these three HLA molecules and T-epitope Designer could be applied for this purpose. Since the frequency of DRB1*0101 and DRB1*0401 alleles of MHC class-II is 20%–50% [29], we selected T-epitopes that interacted with these two HLA molecules during analysis using MHCPred. The final list of epitopes was made with peptide sequences that confirmed the criteria of second level screening (Table 5). Finally, one 9-mer epitope (LRVQWRNNS) from HeV G and two 9-mer epitopes (FRRNNAIAF, LEFRRNNAI) from both NiV M and HeV M were found. However, no efficient T-cell epitope was found from NiV G, NiV F and HeV F protein. Multiple alignment analysis with ClustalW2 (http://www.ebi.ac.uk/Tools/msa/clustalw2/) of 9-mer core epitope (LRVQWRNNS) from G protein showed at least 56% similarity with all HeV G and NiV G proteins retrieved from databases (Figure 1A). On the other hand, the same analysis showed that both 9-mer core epitopes (FRRNNAIAF and LEFRRNNAI) from NiV M and HeV M proteins share 100% identity with all M proteins in HeV and NiV (Figure 1B). For further analysis, we selected two 15-mer epitopes; i) G-epitope (472-VDPLRVQWRNNSSVIS-486) that contained 9 mer core (LRVQWRNNS) and ii) M-epitope (190-PRTMLEFRRNNAIAF-204) that contained both 9 mer core (194-LEFRRNNAI-202 and 196-FRRNNAIAF-204). Physico-chemical properties of G-epitope (472-VDPLRVQWRNNSVIS-486) showed that, it had molecular weight of 1783.0 Da with 253 atoms ($C_{78}H_{127}N_{25}O_{23}$). Analysis with the ProtParam computer program revealed that the half-life of the epitope in mammalian cells was 100 h with instability index 20.38. This result indicated the high stability of the epitope. ProtParam instability index smaller than 40, is predicted as stable. In case of M-epitope (190-PRTMLEFRRNNAIAF-204), analysis with the ProtParam computer program revealed that it was unstable (instability index 50.72). In order to increase stability, we analyzed the instability index of the 11-mer M-epitope (194-LEFRRNNAIAF-204) by ProtParam computer program with different combinations of stable dipeptides. We used two stable dipeptides GK and KG respectively as flanking residues with 194-LEFRRNNAIAF-204 because, Lys (K) and Gly (G) occur with relatively higher frequencies in stable

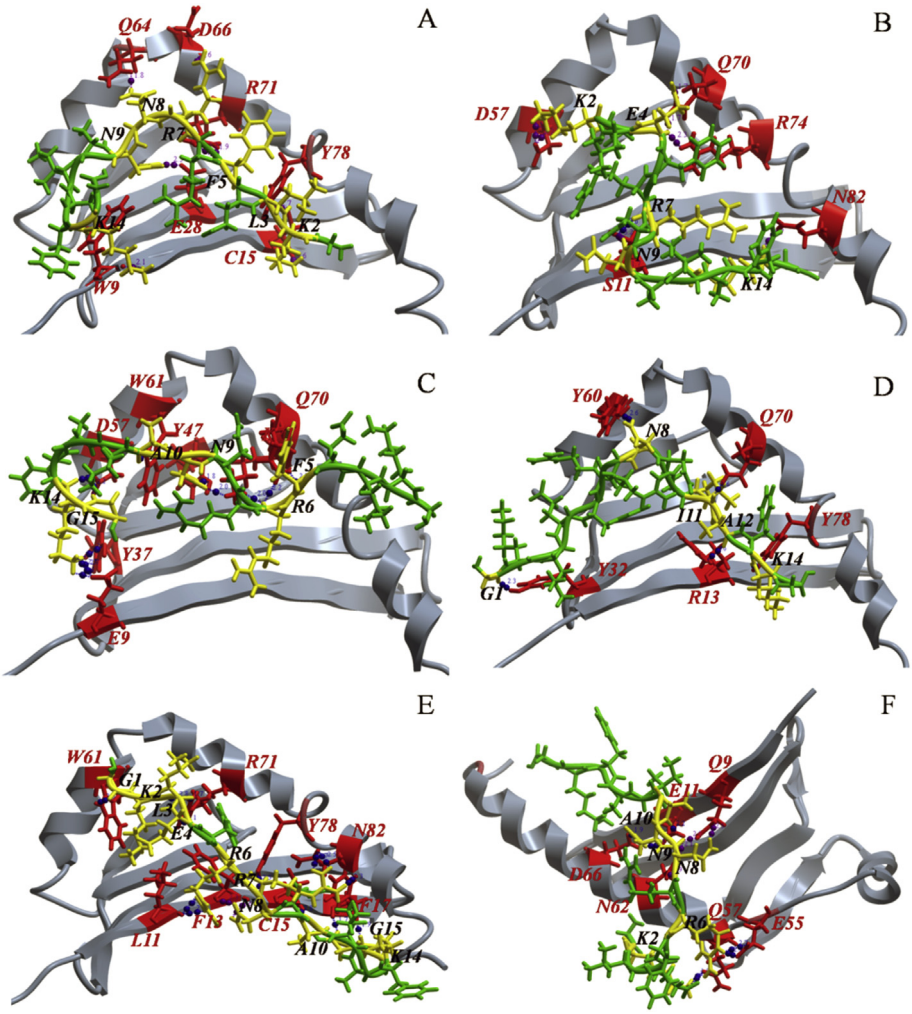

Figure 3. Molecular docking simulation of the M-epitope with MHC II molecules HLA-DRB1*0101 (A), HLA-DRB1*0301 (B), HLA-DRB1*0401 (C), HLA-DRB1*1501 (D), HLA-DRB1*0701 (E) and HLA-DRA (F).
MHC II structures are shown as gray ribbon and the amino acid residues involved in the H-bonding are shown as red sticks and labeled red. The epitope is shown as green ribbon and the amino acid residues involved in the H-bonding are shown as yellow sticks and labeled black. H-bonds are displayed as blue spheres and H-bonding distances are labeled blue.

proteins [30]. Guruprasad *et al*. reported that, dipeptides with Dipeptide Instability Weight Value (DIWV) < 1, may contribute to the stability of the protein [30]. For both dipeptides GK and KG, DIWV is −7.49 according to condition based instability chart of possible 400 dipeptides [30]. ProtParam analysis showed that the stability of the epitope was obtained when dipeptide GK and KG were used as flanking residues (GKLEFRRNNAIAFKG) with the half-life of it 30 h and instability index 28.75. It had molecular weight of 1720.9 Da with 247 atoms ($C_{77}H_{125}N_{25}O_{20}$). Finally, G-epitope (VDPLRVQWRNNSVIS) and M-epitope (GKLEFRRNNAIAFKG) were further analyzed for molecular docking. The 15-mer G-epitope showed at least 66% identity with all NiV and HeV G protein sequences (Figure 1A), while the 15-mer M-epitope with the dipeptide flanking residues showed 73% identity with all NiV and HeV M protein sequences retrieved from the database (Data not shown).

3.4. Molecular docking of predicted epitopes with MHC molecules

The 3D structure of epitope was built by PyMol molecule builder. First, molecular docking was performed to analyze the interaction of 15-mer of the G-epitope (VDPLRVQWRNNSVIS) with MHC II molecules. Docking showed that the protrusion of epitope side chains bound into cavities within the groove of MHC II HLA-DRB1*0101 through 7 hydrogen bonds (Figure 2A). Result revealed that the binding energy between the epitope and the HLA-DRB1*0101 was −216.23 kcal/mol whereas ≤−32 kcal/mol is generally considered as good score in ICM method [31]. As HLA-DRA does not have polymorphisms in the peptide binding part and acts as the sole alpha chain for HLA-DRB1, HLA-DRB3, HLA-DRB4 and HLA-DRB5, molecular docking simulation between the G-epitope and HLA-DRA was performed (Figure 2F). Docking result revealed that the G-epitope was bound to the groove by 7 hydrogen bonds. The binding energy was −214.55 kcal/mol. Molecular docking of the epitope was further performed with HLA-DRB1*0301, HLA-DRB1*0401, HLA-DRB1*1501 and HLA-DRB1*0701 (Figure 2). Docking study showed that the epitope side chains bound into cavities within the grooves of HLA-DRB1*0301, HLA-DRB1*0401, HLA-DRB1*1501 and HLA-DRB1*0701 through 12, 8, 12 and 7 hydrogen bonds, respectively, and the binding energy of the epitope with these MHC II molecules were −229.53, −223.02, −259.64 and −231.49 kcal/mol, correspondingly.

Similarly, molecular docking of M-epitope (GKLEFRRNNAIAFKG) with HLA-DRB1*0101 showed that the epitope was bound to the groove by 8 hydrogen bonds (Figure 3A) with −224.91 kcal/mol binding energy. Analysis of molecular docking simulation between the epitope and HLA-DRA (Figure 3F) revealed that the epitope was bound to the groove by 14 hydrogen bonds with −201.53 kcal/mol binding energy. Molecular docking result of the epitope with HLA-DRB1*0301, HLA-DRB1*0401, HLA-DRB1*1501 and HLA-DRB1*0701 (Figure 3) revealed that the epitope side chains bound into cavities within the grooves of these MHC II molecules through 8, 16, 6 and 15 hydrogen bonds, respectively. The binding energy was −189.33, −234.24, −229.14 and −262.71 kcal/mol, correspondingly. Docking Results indicated that the both predicted epitope could bind with MHC II molecules in antigen presenting cells and might trigger both B-cell and T-Cell mediated immunity since epitope binding within the grove of the MHC II molecules is considered as principal determinant of binding affinity [21,22].

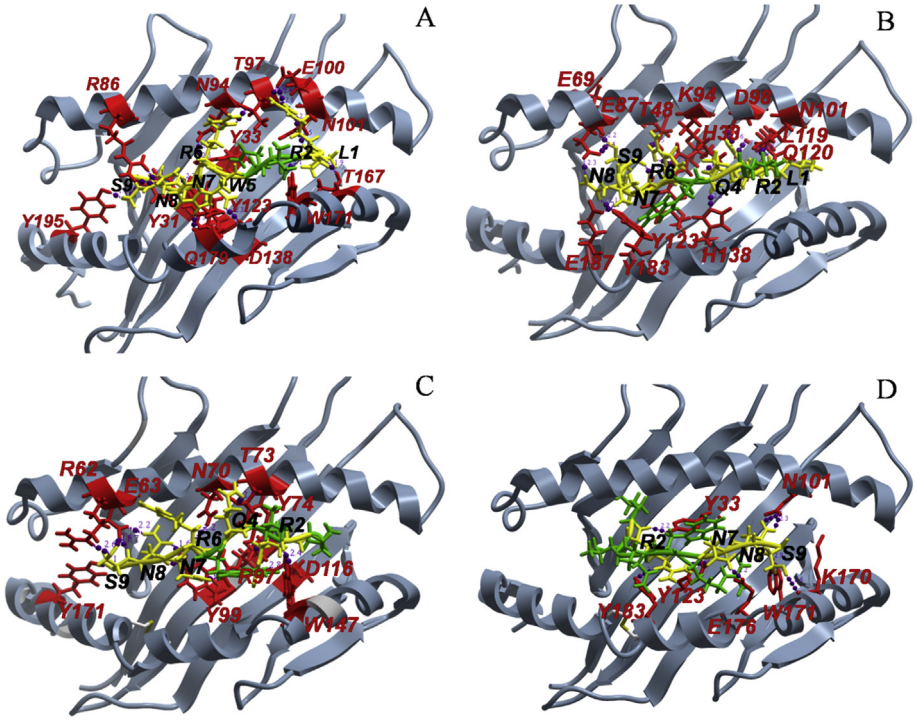

Figure 4. Molecular docking simulation of the 9-mer LRVQWRNNS from G-epitope with HLA-B7 (A), HLA-B27 (B), HLA-B44 (C) and HLA-B62 (D). MHC II structures are shown as gray ribbon and the amino acid residues involved in the H-bonding are shown as red sticks and labeled red. The epitope is shown as green ribbon and the amino acid residues involved in the H-bonding are shown as yellow sticks and labeled black. H-bonds are displayed as blue spheres and H-bonding distances are labeled blue.

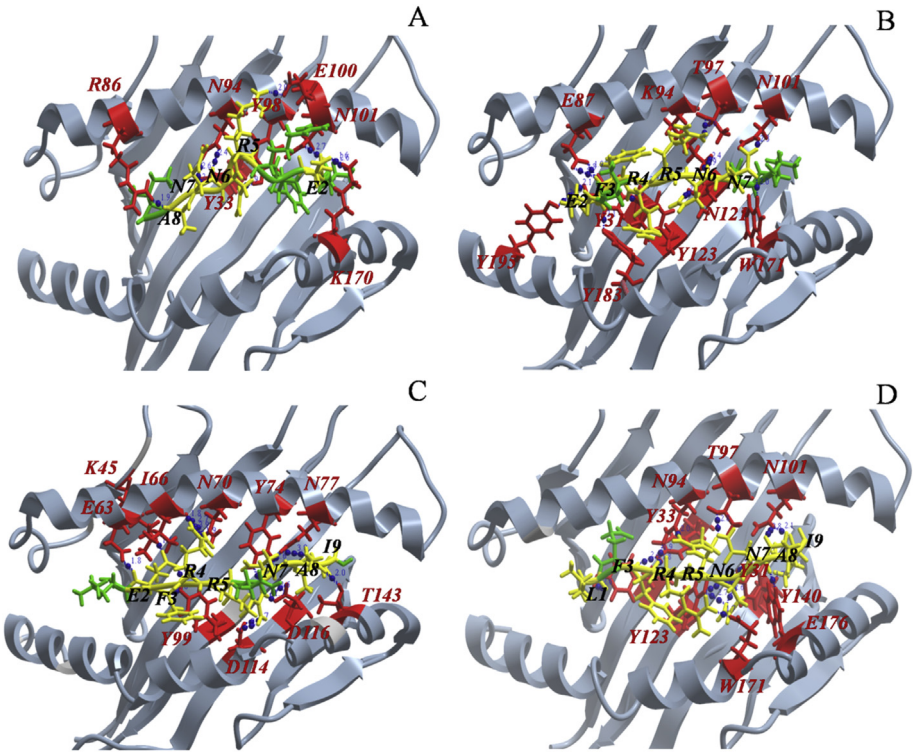

Figure 5. Molecular docking simulation of the 9-mer LEFRRNNAI from M-epitope with HLA-B7 (A), HLA-B27 (B), HLA-B44 (C) and HLA-B62 (D). MHC II structures are shown as gray ribbon and the amino acid residues involved in the H-bonding are shown as red sticks and labeled red. The epitope is shown as green ribbon and the amino acid residues involved in the H-bonding are shown as yellow sticks and labeled black. H-bonds are displayed as blue spheres and H-bonding distances are labeled blue.

Docking study of the 9 mer of the predicted G-epitope with MHC I molecule was also performed. The analysis showed that the LRVQWRNNS bound into cavities within the groove of HLA-B7, HLA-B27, HLA-B44, HLA-B62 through 17, 19, 13, 8 hydrogen bonds respectively, and the binding energy with these MHC I molecules were −242.61, −184.91, −210.34, −232.55 kcal/mol, correspondingly (Figure 4). Another docking study of the 9 mer of the predicted M-epitope with MHC I molecule showed that the LEFRRNNAI bound into cavities within the groove of HLA-B7, HLA-B27, HLA-B44, HLA-B62 through 10, 16, 22, 17 hydrogen bonds respectively, and the binding energy with these MHC I molecules were −216.26, −200.18, −198.30 and −245.28 kcal/mol, correspondingly (Figure 5). The binding energy of the 9 mer of the predicted G-epitope and M-epitope with MHC I molecule might be significant since binding energy below −32 kcal/mol is biologically considerable [21,22]. Multiple alignment analysis of MHC II molecules and HLA-DRA showed that these molecules have difference in sequence and as well as alignment of HLA super types B7, B27, B44 and B62 showed the dissimilarity of their sequence (data not shown).

4. Discussion

NiV and HeV emerged in the mid to late 1990s causing disease outbreaks in livestock and people. HeV appeared first in Queensland, Australia in 1994 causing a severe respiratory disease in horses along with a human case fatality. And, NiV emerged a few years later in Malaysia and Singapore in 1998–1999 causing a large outbreak of encephalitis with high mortality in people and also a respiratory disease in pigs. In people, both HeV and NiV are also capable of causing relapsed encephalitis. *Henipaviruses* are becoming deadly day by day. Spillovers of HeV into horses continue to occur in Australia and NiV has caused outbreaks in people in Bangladesh and India nearly annually since 2001, making these zoonotic pathogens important transboundary biological threats [1]. Therefore, it is important to take preventive measures against *Henipavirus*. Ribavirin used alone and in combination with Chloroquine previously against NiV and HeV, did not show much success [8–10]. To date, there is no recommended vaccine for *Henipaviruses* which is necessary to prevent. The concept of prevention of viruses by designing a common, epitope-based vaccine has not been reported previously and epitope-based vaccine offers the ability to immunize with a minimal structure and it will stimulate an effective specific immune response, while avoiding potential undesirable effects.

In this study, we aimed to design an epitope-based common vaccine for both NiV and HeV. For this purpose we analyzed *Henipavirus* membrane proteins which are F, G and M proteins. Prediction of antigenicity of the non-identical sequences from membrane proteins were done using VaxiJen server [12] and TMHMM [14] prediction server used to identify exo-membrane amino acid sequences of each protein. We predicted antigenic B-cell epitopes using both BCpred and AAP prediction methods [15], which were finally used for prediction of T-cell epitopes. From the analysis, we predicted one 9-mer epitope (LRVQWRNNS) from HeV G and two 9-mer epitopes (FRRNNAIAF, LEFRRNNAI) from both NiV M and HeV M. We also checked similarities of these three epitopes individually with all Hev and NiV corresponding proteins which showed at least 56% identity. For our further study we selected two 15-mer

epitopes; i) G-epitope (472- VDPLRVQWRNNSSVIS-486) that contained 9 mer core (LRVQWRNNS) and ii) M-epitope (190-PRTMLEFRRNNAIAF-204) that contained both 9 mer core (194-LEFRRNNAI-202 and 196-FRRNNAIAF-204). To predict if these epitopes are poorly immunogenic or unstable, physicochemical properties of both epitopes were checked with ProtParam program. ProtParam computer program revealed that unlike G-epitope, M-epitope (190-PRTMLEFRRNNAIAF-204) was unstable. The stability of the epitope was obtained when dipeptide were used as flanking residues (GKLEFRRNNAIAFKG) [30]. With a view to designing common epitope-based vaccine for both NiV and HeV, we analyzed the identity of predicted G- and M-epitopes with corresponding NiV and HeV protein sequences retrieved from the database. G-epitope shares at least 66% identity while M-epitope with the dipeptide flanking residues shares 73% identity. Finally, both G-epitope and M-epitope were subjected to molecular docking. The molecular docking result with both MHC I and II molecules were significant with biologically considerable binding energy.

These epitopes are found to be stable and meet crucial requirements for an epitope to be used as a vaccine. This reflects a promising scope to use these epitopes as common preventive and therapeutic treatment for NiV and HeV. As these epitopes remain long been similar and conserved since 1994, it may be possible to use in future. Though this epitope-based vaccine is designed by *in silico* analyses, the actual immunogenicity, stability, efficacy and their delivery strategy inside the recipient's body can not be determined by this *in silico* analysis. To address these questions *in vitro* and *in vivo* experiments are essential.

In the present study, we conducted analysis of F, G and M proteins from HeV and NiV by using different bioinformatics tools to identify the best antigen. From the databases (NCBI Protein database, UNIPROT KB), we retrieved all the non-identical sequences of F (11 sequences), G (14 sequences) and M (11 sequences) proteins of both HeV and NiV. Following analysis with various bioinformatics tools, two 15-mer epitopes, G-epitope (VDPLRVQWRNNSVIS) from G proteins of HeV and NiV and M-epitope (GKLEFRRNNAIAFKG) containing the 11-mer core epitope (LEFRRNNAIAF) from both NiV and HeV M proteins have been identified and characterized as novel efficient B- and T-cell epitope. Both of these epitopes showed interaction with at least six MHC II molecules with significant numbers of hydrogen bonding, which are of immense importance in case of inducing strong immunity. The 9 mer of both G-epitope and M-epitope also showed interaction with MHC I molecules with significant hydrogen bonding. Data in our present study revealed the notion that the epitopes from G and M protein might be the target for peptide-based subunit vaccine design against HeV and NiV. However, biochemical analysis is necessary to experimentally validate the interaction of epitopes individually with the MHC molecules through elucidation of immunity induction.

Conflict of interest statement

Authors disclose no potential conflict of interests.

References

[1] Solomon T, Hart IJ, Beeching NJ. Viral encephalitis: a clinician's guide. *Pract Neurol* 2007; **7**(5): 288-305.

[2] Islam MS, Sazzad HM, Satter SM, Sultana S, Hossain MJ, Hasan M, et al. Nipah virus transmission from bats to humans associated with drinking traditional liquor made from Date Palm Sap, Bangladesh, 2011-2014. *Emerg Infect Dis* 2016; **22**(4): 664-670.

[3] Clayton BA. Nipah virus: transmission of a zoonotic paramyxovirus. *Curr Opin Virol* 2017; **22**: 97-104.

[4] Mire CE, Satterfield BA, Geisbert JB, Agans KN, Borisevich V, Yan L, et al. Pathogenic differences between Nipah virus Bangladesh and Malaysia strains in primates: implications for antibody therapy. *Sci Rep* 2016; **6**: 30916.

[5] Marsh GA, Wang L-F. Hendra and Nipah viruses: why are they so deadly? *Curr Opin Virol* 2012; **2**(3): 242-247.

[6] Choi C. Nipah's return. The lethal “flying fox” virus may spread between people. *Sci Am* 2004; **291**(3): 21A, 22.

[7] Wang L-F, Yu M, Hansson E, Pritchard LI, Shiell B, Michalski WP, et al. The exceptionally large genome of Hendra virus: support for creation of a new genus within the family Paramyxoviridae. *J Virol* 2000; **74**(21): 9972-9979.

[8] Chong HT, Kamarulzaman A, Tan CT, Goh KJ, Thayaparan T, Kunjapan SR, et al. Treatment of acute Nipah encephalitis with ribavirin. *Ann Neurol* 2001; **49**(6): 810-813.

[9] Playford EG, McCall B, Smith G, Slinko V, Allen G, Smith I, et al. Human Hendra virus encephalitis associated with equine outbreak, Australia, 2008. *Emerg Infect Dis* 2010; **16**(2): 219.

[10] Steffen DL, Xu K, Nikolov DB, Broder CC. Henipavirus mediated membrane fusion, virus entry and targeted therapeutics. *Viruses* 2012; **4**(2): 280-308.

[11] Vivona S, Gardy JL, Ramachandran S, Brinkman FS, Raghava G, Flower DR, et al. Computer-aided biotechnology: from immuno-informatics to reverse vaccinology. *Trends Biotechnol* 2008; **26**(4): 190-200.

[12] Doytchinova IA, Flower DR. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinform* 2007; **8**(1): 4.

[13] Prabhavathy K, Perumal P, Sundara Baalaji N. In silico identification of B- and T-cell epitopes on OMPLA and LsrC from *Salmonella typhi* for peptide-based subunit vaccine design. *Indian J Biotechnol* 2011; **10**(4): 440-451.

[14] Krogh A, Larsson B, Von Heijne G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* 2001; **305**(3): 567-580.

[15] EL-Manzalawy Y, Dobbs D, Honavar V. Predicting linear B-cell epitopes using string kernels. *J Mol Recognit* 2008; **21**(4): 243-255.

[16] Singh H, Raghava G. ProPred1: prediction of promiscuous MHC Class-I binding sites. *Bioinformatics* 2003; **19**(8): 1009-1014.

[17] Singh H, Raghava G. ProPred: prediction of HLA-DR binding sites. *Bioinformatics* 2001; **17**(12): 1236-1237.

[18] Sturniolo T, Bono E, Ding J, Raddrizzani L, Tuereci O, Sahin U, et al. Generation of tissue-specific and promiscuous HLA ligand databases using DNA microarrays and virtual HLA class II matrices. *Nat Biotechnol* 1999; **17**(6): 555-561.

[19] Guan P, Doytchinova IA, Zygouri C, Flower DR. MHCPred: a server for quantitative prediction of peptide–MHC binding. *Nucleic Acids Res* 2003; **31**(13): 3621-3624.

[20] Kangueane P, Sakharkar MK. T-Epitope designer: a HLA-peptide binding prediction server. *Bioinformation* 2005; **1**(1): 21.

[21] Patronov A, Dimitrov I, Flower DR, Doytchinova I. Peptide binding prediction for the human class II MHC allele HLA-DP2: a molecular docking approach. *BMC Struct Biol* 2011; **11**(1): 32.

[22] Totrov M, Abagyan R. Flexible protein-ligand docking by global energy optimization in internal coordinates. *Proteins* 1997; **29**(s1): 215-220.

[23] Schwede T, Kopp J, Guex N, Peitsch MC. SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Res* 2003; **31**(13): 3381-3385.

[24] Laskowski RA. PDBsum: summaries and analyses of PDB structures. *Nucleic Acids Res* 2001; **29**(1): 221-222.

[25] Azad AK, Hasan MM, Hossain MS, Rahman MR, Chowdhury PA. *In silico* analysis of outer membrane protein 31 of *Brucella* spp. to

identify and characterize the potential t cell epitope. *Int J Pharm Med Bio Sci* 2013; **2**(3): 28-45.

[26] Hsu VP. Nipah and Hendra viruses. *Persp Med Virol* 2006; **16**: 179.

[27] Wang L-F, Harcourt BH, Yu M, Tamin A, Rota PA, Bellini WJ, et al. Molecular biology of Hendra and Nipah viruses. *Microbes Infect* 2001; **3**(4): 279-287.

[28] Somvanshi P, Seth P. Prediction of T cell epitopes for the utility of vaccine development from the structural proteins of dengue virus variants using *in silico* methods. *Indian J Biotechnol* 2009; **8**(2): 193-198.

[29] Panigada M, Sturniolo T, Besozzi G, Boccieri MG, Sinigaglia F, Grassi GG, et al. Identification of a promiscuous T-cell epitope in *Mycobacterium tuberculosis* Mce proteins. *Infect Immun* 2002; **70**(1): 79-85.

[30] Guruprasad K, Reddy BB, Pandit MW. Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting *in vivo* stability of a protein from its primary sequence. *Protein Eng* 1990; **4**(2): 155-161.

[31] Khan JM, Ranganathan S. pDOCK: a new technique for rapid and accurate docking of peptide ligands to major histocompatibility complexes. *Immunome Res* 2010; **6**(Suppl 1): S2.