



# Immunoinformatics Approach to Design T-cell Epitope-Based Vaccine Against Hendra Virus

Mohit Kamthania<sup>1,2</sup> · Sukrit Srivastava<sup>2</sup> · Meha Desai<sup>3</sup> · Anubhav Jain<sup>1</sup> · Archana Shrivastav<sup>4</sup> · D. K. Sharma<sup>1</sup>

Accepted: 29 December 2018 / Published online: 9 January 2019  
© Springer Nature B.V. 2019

## Abstract

Screening of HLA class II epitope-based peptides as potential vaccine candidates is one of the most rational approach for vaccine development against Hendra virus (HeV) infection, for which currently there is no successful vaccine in practice. In this study, screening of epitopes from HeV proteins viz matrix, glycoprotein, nucleocapsid, fusion, C protein, V protein, W protein and polymerase, followed by highest binding affinity & molecular dynamic simulation of selected T-cell epitopes with their corresponding HLA class II alleles has been done. The server ProPred facilitates the binding prediction of HLA class II allele specific epitopes from the antigenic protein sequences of HeV. PEPstrMOD server was used for PDB structure modeling of the screened epitopes and MODELLER was used for HLA alleles modeling. We docked the selected T-cell epitopes with their corresponding HLA allele structures using the AutoDock 4.2 tool. Further the selected docked complex structures were optimized by Nanoscale Molecular Dynamics program (NAMD) at 5 ps, with the CHARMM-22 force field parameter incorporated in Visual Molecular Dynamics (VMD 1.9.2) and complex structure stability was evaluated by calculating RMSD values. Epitopes IRIFVPATN (Nucleocapsid), MRNLLSQL (Nucleocapsid), VRRAGKYYS (Matrix) and VRLKCLLCG (Fusion) proteins have shown considerable binding with DRB1\*0806, DRB1\*1304, DRB1\*0701 and DRB1\*0301 HLA class II allele respectively. Toxicity, antigenicity and population coverage of epitopes IRIFVPATN, MRNLLSQL, VRRAGKYYS and VRLKCLLCG were analyzed by Toxin Pred, Vexijen and IEDB tool, respectively. The potential T-cell epitopes can be utilized in designing comprehensive epitope-based vaccines and diagnostic kits against Hendra virus after further in-vivo studies.

**Keywords** Hendra virus · T-cell epitope · Antigenicity · HLA alleles, Modeling · Population coverage · Vaccine designing

## Introduction

Hendra virus (HeV) initially emerged in an outbreak of acute respiratory disease amongst horses and humans in Australia, 1994 (Murray et al. 1995). Patients of these cases were found to have infected with close contact to diseased horses. Initial cases of Hendra infection clinically diagnosed as severe respiratory infection (Selvey et al. 1995; Hanna et al. 2006; O'Sullivan et al. 1997; Playford et al. 2010). There is no

specific treatment for human cases of Hendra virus. Flying foxes (pteropterid bats) considered as the natural reservoir of HeV. Transmission of viral infection to humans occurred likely due to exposure of mucous membrane or non-intact skin to nasal secretions, urine and blood of infected horses (Mire et al. 2015). HeV was primarily isolated from infected uterine fluid and fetal tissue of bat species *Pteropus poliocephalus* and *Pteropus alecto* respectively (Halpin et al. 2000). Equivac Hev, a subunit vaccine from glycoprotein of HeV is the only approved licensed vaccine from Australian government which is used for prevention of Hev in the horse population.

Peptide vaccines are considered an alternative to classical vaccines that are trying to address issues of possible vaccine side effects related to vaccination with a heterogeneous multicomponent preparation. The peptide-based vaccines include chemical approach to synthesize the selected epitopes that are specific and trigger immune responses.

✉ D. K. Sharma  
sharmadr\_dinesh@yahoo.in

<sup>1</sup> Jiwaji University, Gwalior, Madhya Pradesh, India

<sup>2</sup> Mangalayatan University, Aligarh, Uttar Pradesh, India

<sup>3</sup> Veer Narmad South Gujarat University, Surat, Gujarat, India

<sup>4</sup> Dept of Microbiology, College of Life Science, Cancer Hospital Campus Gwalior, Gwalior, Madhya Pradesh, India

**Table 1** Template PDB ID for modeling of selected HLA class II alleles

S. no.	HLA class II allele	PDB ID of template
1	DRB1*0817	1A6A
2	DRB1*1101	1YMM
3	DRB1*0405	2SEB
4	DRB1*0410	2SEB
5	DRB1*1104	2WBJ
6	DRB1*0301	1A6A
7	DRB1*1301	2WBJ
8	DRB1*0701	1AQD
9	DRB1*1102	2WBJ
10	DRB1*1128	1A6A
11	DRB1*0402	4MDI
12	DRB1*1304	1YMM
13	DRB1*0804	2SEB
14	DRB1*0806	2WBJ

Immunoinformatic focuses mainly on prediction of potential epitopes which brings down the laboratory analysis cost & time of potential vaccine candidate. Using immunoinformatic tools, immunologist can screen and analyze short sequences from the full-length foreign proteins which can act as an immunogenic epitopes & facilitating to be a vaccine candidate (Li et al. 2014; Kamthania and Sharma 2016). Important implication of this study is to screen promiscuous T-cell epitopes from HeV proteins viz matrix, glycoprotein, nucleocapsid, fusion, C protein, V protein, W protein, and polymerase. These predicted promiscuous T-cell epitopes may be the promising targets for epitope-based vaccine design for HeV.

**Table 2** IC<sub>50</sub> values of alleles with their corresponding predicted potential T-cell epitopes IRIFVPATN, MRNLLSQSL, VRRAGKYYS and VRLKCLLCG with an affinity of < 500 nM

T-cell epitope	Total no. of HLA-peptide binders	MHC-I alleles with IC <sub>50</sub> values (bracketed)	MHC-II alleles with IC <sub>50</sub> values (bracketed)
IRIFVPATN	6	HLA-A*02:03(16.00), HLA-A*02:06(439.54), HLA-A*03:01(334.20), HLA-A*11:01(34.04), HLA-A*68:01(59.16), HLA-A*68:02(192.75)	–
MRNLLSQSL	9	HLA-A*02:03(11.72), HLA-A*02:06(510.50), HLA-A*03:01(340.41), HLA-A*11:01(31.84), HLA-A*31:01(411.15), HLA-A*68:01(180.72), HLA-B*35:01(408.32)	HLA-DRB*01:01(35.97), HLA-DRB*07:01(428.55)
VRRAGKYYS	7	HLA-A*02:02(179.47), HLA-A*02:06(193.20), HLA-A*11:01(75.68), HLA-A*68:01(457.09)	HLA-DRB*01:01(381.94), HLA-DRB*04:01(210.38), HLA-DRB*07:01(297.17)
VRLKCLLCG	6	HLA-A*02:01(112.46), HLA-A*02:03(151.01), HLA-A*02:06(57.94), HLA-A*11:01(123.88), HLA-A*68:01(382.82)	HLA-DRB*04:01(51.29)

## Materials and Methods

### Retrieval of Amino Acid Sequence

The amino acid sequences of matrix, glycoprotein, nucleocapsid, fusion, C protein, V protein, W protein and polymerase, were retrieved from NCBI protein sequence database (<http://www.ncbi.nlm.nih.gov/protein>). The total of 22 nucleocapsid, 20 matrix, 22 fusion, 22 glycoprotein, 14 W protein, 17 V protein, 15 C protein and 15 polymerase protein sequences from different HeV strains, available at NCBI database were retrieved & downloaded in FASTA format.

### Antigenicity Prediction

VaxiJen server (<http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html>) (Doytchinova and Flower 2007) was used with default parameters to predict the antigenicity of HeV candidate proteins. The threshold value of vaxiJen was set to 0.4 (default for viruses). The viral proteins having VaxiJen score above 0.4 values were considered to be antigenic in nature. Antigenic proteins were chosen for further study.

### Promiscuous T-Cell Epitope Prediction

The prediction of promiscuous T-cell epitope binder to HLA class-II alleles was performed by ProPred (Singh and Raghava 2001). ProPred is an on-line web tool; it utilizes matrix-based prediction of HLA (Class II alleles) binding sites in an antigenic protein sequence.

### 3D Structure Prediction of Promiscuous T-cell Epitopes

The PEPstrMOD (Singh et al. 2015) method performed to find out the tertiary structure of selected small epitope with

**Table 3** Conservancy analysis using Conservation across antigen tool of IEDB revealed that all the four chosen epitopes are 100% conserved across all the HeV strain's respective protein sequences retrieved from NCBI database

S. no.	Protein name	Amino acid sequence of selected epitope	Selected epitope length	Position	Percent of protein sequence match at identity $\geq$ 100%
1	Nucleocapsid	IRIFVPATN	9	35–43	100.00% (22/22)
2	Nucleocapsid	MRNLLSQQSL	9	217–225	100.00% (22/22)
3	Matrix	VRRAGKYYS	9	243–251	100.00% (20/20)
4	Fusion	VRKCLLCG	9	6–14	100.00% (22/22)

**Table 4** VaxiJen Results of antigenicity

Protein	Overall antigen prediction
Nucleocapsid	0.5385 (Probable ANTIGEN)
Matrix	0.4422 (Probable ANTIGEN)
Fusion	0.5529 (Probable ANTIGEN)
Glycoprotein	0.5380 (Probable ANTIGEN)
W protein	0.4747 (Probable ANTIGEN)
nonstructural protein V	0.4978 (Probable ANTIGEN)
nonstructural protein C	0.3540 (Probable NON-ANTIGEN)
Polymerase	0.4843 (Probable ANTIGEN)

sequence length of nine residues. The PEPstrMOD tool prediction strategy utilizes the secondary structure data anticipated by PSIPRED (Kaur and Raghava 2004) and  $\beta$ -turns data anticipated by BetaTurns (Jones et al. 1999).

### Modeling of HLA Class II Alleles

The sequences of alleles viz DRB1\*1304, DRB1\*0804, DRB1\*0405, DRB1\*0806, DRB1\*0402, DRB1\*0701, DRB1\*1301, DRB1\*1104, DRB1\*1102, DRB1\*0410, DRB1\*1128, DRB1\*1101, DRB1\*0301, & DRB1\*0817 were downloaded from IMGT/HLA database (<http://www.ebi.ac.uk/ipd/imgt/hla/allele.html>) (Robinson et al. 2012). BLASTP program (Altschul et al. 1990) was used to find the templates for these allele sequences and respective PDB IDs (Table 1) retrieved from Protein Data Bank (<http://www.rcsb.org/pdb>). MODELLER 9.17 (Šali et al. 1995) was used for homology modeling of HLA alleles. Five models were generated by modeller software and the best model chosen based on their lowest predicted discrete optimized protein energy (DOPE) score. The modeled alleles overall quality was determined using Ramachandran plot analysis utilizing PROCHECK (Laskowski et al. 1993), which detect the geometrical orientation of each amino acid residues with respect to stereochemical parameters.

### Molecular Docking

AutoDock 4.2 (Morris et al. 1998) was utilized to form the docking complex of predicted T-cell epitopes and HLA

alleles. In case of rigid protein & flexible ligand during auto-dock, the Lamarckian Genetic Algorithm (LGA) was used. The best conformation of docked complex was chosen on the basis of minimum binding energy and best fitting of epitope-HLA allele complex with highest number of H-bonds formation. Python Molecular Viewer (Sanner et al. 1999) used to visualize the docked complex of HLA allele with predicted T cell epitope.

### Docked Complex Stability Validation by MD Simulation

NAMD used for MD simulation of selected docked complex (Phillips et al. 2005). VMD analyses & view the results of MD simulation & also interface with NAMD (Humphrey et al. 1996). Protein structure file (psf) was made using the topology files and initial pdb files of the HLA II allele-epitope docked complex utilizing psfgen package of VMD. NAMD created the trajectory DCD file. The result of simulation was analyzed by calculation the Root mean square deviation (RMSD) of the docked complex. The rmsd.dat file contains the value of RMSD which was further analyzed by Microsoft office excel.

### Toxicity Prediction of the Selected T-Cell Epitopes

ToxinPred (<http://crdd.osdd.net/raghava/toxinpred/>) (Gupta et al. 2013) was utilized to calculate toxicity of predicted T- cell epitopes. ToxinPred is an in-silico method to predict toxic/non-toxic peptides. ToxinPred was run with default parameters and only non-toxic T-cell epitopes were selected for further study.

### HLA-Distribution Analysis

MHCPred (<http://www.ddg-pharmfac.net/mhcpred/MHCPred/>) was utilized to select the high affinity HLA class II binder for selected epitopes with their  $IC_{50}$  (half maximal inhibitory concentration) value. The alleles having cut-off value of  $IC_{50}$  between 0.01 and 500 nM were selected.

**Table 5** ProPred predicted T-cell epitope for HLA Class II with binding scores

Protein name	Accession no. (NCBI)	Amino acid length	Start position	Epitopes	HLA class II alleles	Propred (% of highest score)
Nucleocapsid	NP_047106	532	128	MRILKTARE	DRB1*0410	84.04
					DRB1*0806	75.58
					DRB1*0405	73.4
			34	IRIFVPATN	DRB1*1304	68.89
					DRB1*1304	62.22
					DRB1*0806	61.63
					DRB1*0817	55.45
Matrix	NP_047110	352	242	VRRAGKYYS	DRB1*0701	51.72
					DRB1*1301	57.95
			261	MKLQFSLGS	DRB1*0301	55.26
					DRB1*0421	52.22
					DRB1*1104	51.81
Fusion	NP_047111	546	493	MIILYVLSI	DRB1*1104	73.49
					DRB1*0701	54.31
			315	VLIRNTLIS	DRB1*0701	54.31
					DRB1*0402	73.96
Glycoprotein	NP_047112	604	5	VRLKCLLCG	DRB1*0806	66.28
					DRB1*0701	68.1
			314	FVLIRNTLI	DRB1*0701	68.1
					DRB1*0301	63.16
					DRB1*1301	60.23
W protein	YP_007188593	447	355	LRLREPPQS	DRB1*1102	60.71
					DRB1*0806	63.95
					DRB1*0804	68.75
V protein	NP_047108	457	355	LRLREPPQS	DRB1*0806	63.95
					DRB1*0806	63.95
			304	VRRKDSLMLQ	DRB1*0806	70.93
					DRB1*0817	68.32
					DRB1*0804	66.25
Polymerase	NP_047113	2244	1169	FLILNRLLS	DRB1*1305	82.76
					DRB1*1128	82.76
					DRB1*1101	81.93
			468	MYMKDKALS	DRB1*0804	82.5
					DRB1*0806	76.74

## Population Coverage Analysis

IEDB (Immune Epitope Database and Analysis Resource) population coverage tool ([http://tools.immuneepitope.org/tools/population/iedb\\_input](http://tools.immuneepitope.org/tools/population/iedb_input)) (Bui et al. 2006) was used to study the worldwide geographical population coverage by the selected epitopes and HLA Class II alleles pair as resulted from MHCpred (Table 2). Default parameters were set while running the tool. IEDB conservancy tool consist data set frequencies of 3245 alleles for 115 countries, 21 ethnicities and 16 geographical areas. The prediction was taken utilizing the most resent information to set from Allele Frequency Net Database (Gonzalez-Galarza et al. 2010) at IEDB.

## Conservancy Analysis of Selected Epitopes

To find out the degree of conservation, all the selected epitopes were aligned against all respective source protein sequences retrieved from NCBI database using EBI-clustal Omega program (Sievers et al. 2011). The Multiple sequence alignment (MSA) was visualized using Jalview (Waterhouse et al. 2009). Conservancy analysis was again performed for the selected T-cell epitopes by conservancy tool of IEDB (Table 3) (Bui et al. 2007).

**Table 6** Toxicity prediction of the peptides by ToxinPred

Peptide sequence	SVM Score	Prediction	Hydrophobicity	Hydropathicity	Hydrophilicity	Charge	Mol. wt
MRILKTARE	-1.61	Non-toxin	-0.41	-0.57	0.69	2.00	1117.50
IRIFVPATN	-1.23	Non-toxin	0.02	0.83	-0.59	1.00	1030.36
LNLRSRLAA	-1.09	Non-toxin	-0.26	0.19	0.01	2.00	1013.33
MRNLLSQSL	-1.32	Non-toxin	-0.20	0.02	-0.30	1.00	1061.40
VRRAGKYYS	-0.95	Non-toxin	-0.43	-1.19	0.30	3.00	1099.37
MKLQFSLGS	-0.98	Non-toxin	-0.02	0.32	-0.40	1.00	1010.35
IKINGVISK	-1.02	Non-toxin	-0.02	0.58	-0.04	2.00	0971.35
FRRNNAIAF	-0.97	Non-toxin	-0.26	-0.26	-0.16	2.00	1108.38
MIILVLSI	-0.99	Non-toxin	0.42	2.79	-1.53	0.00	1064.54
VLIRNTLIS	-0.77	Non-toxin	0.02	1.26	-0.62	1.00	1028.40
VRLKCLLCG	-0.69	Non-toxin	-0.05	1.31	-0.32	2.00	1004.45
FVLIRNTLI	-0.74	Non-toxin	0.12	1.66	-0.93	1.00	1088.50
IRLAVRPKS	-1.22	Non-toxin	-0.32	-0.11	0.41	3.00	1039.41
LKIHECNIS	-0.64	Non-toxin	-0.11	0.04	-0.04	0.50	1056.38
LRVQWRNNS	-0.61	Non-toxin	-0.48	-1.47	0.02	2.00	1172.44
LRLREPPQS	-0.72	Non-toxin	-0.46	-1.38	0.66	1.00	1095.39
LRLREPPQS	-0.72	Non-toxin	-0.46	-1.38	0.66	1.00	1095.39
VRRKDSLMLQ	-1.26	Non-toxin	-0.55	-1.20	0.88	2.00	1132.47
LRMMEVLKE	-1.47	Non-toxin	-0.22	0.02	0.48	0.00	1148.58
LQFLKKLKLGK	-0.70	Non-toxin	-0.18	-0.16	0.14	3.00	1074.52
FLILNRLLS	-1.28	Non-toxin	0.09	1.52	-0.89	1.00	1088.50
MYMKDKALS	-0.96	Non-toxin	-0.21	-0.44	0.23	1.00	1086.45

## Results and Discussion

### Antigenicity Prediction

Amino acid sequences of matrix, glycoprotein, nucleocapsid, fusion, C protein, V protein, W protein and polymerase proteins were screened by VaxiJen. All the proteins were found antigenic except one nonstructural protein C which is nonantigenic at threshold value 0.4 (default threshold for viral proteins) (Table 4). Fusion protein had been predicted as highest antigenic with score of 0.5529 among all candidate proteins.

### Promiscuous T-cell Epitope selection and analysis

Antigenic HeV proteins were subjected to ProPred for selection of HLA Class II specific T-cell epitopes binders. Epitopes showing highest score with the maximum number of HLA Class II alleles binders were selected at a threshold value of 4% (Table 5).

### Toxicity Prediction of the Peptide Epitopes

ToxinPred (Gupta et al. 2013) used for toxicity prediction of selected T-cell epitopes. ToxinPred tool is a unique

in-silico method based on Support Vector Machine (SVM) in predicting toxicity of peptides along with important physico-chemical properties viz hydropathicity, hydrophilicity, hydrophobicity, charge and molecular weight. The selected epitopes were subjected to ToxinPred and only non-toxic T-cell epitopes were selected for further studies (Table 6).

### Structure Modeling of T-cell Epitopes and Alleles

3D structures of selected epitopes were predicted by PEPstrMOD. Modeller 9.17 was employed to generate homology model of alleles: DRB1\*1304, DRB1\*0804, DRB1\*0405, DRB1\*0806, DRB1\*0402, DRB1\*0701, DRB1\*1301, DRB1\*1104, DRB1\*1102, DRB1\*0410, DRB1\*1128, DRB1\*1101, DRB1\*0301, & DRB1\*0817; template used (PDB ID): 1YMM, 2SEB, 2SEB, 2WBJ, 4MDI, 1AQD, 2WBJ, 2WBJ, 2WBJ, 2SEB, 1A6A, 1YMM, 1A6A & 1A6A respectively (Table 1). PROCHECK was utilized to analyze selected models quality. Ramachandran result of HLA alleles model (DRB1\*0701, DRB1\*0301, DRB1\*1304 and DRB1\*0806 alleles) forming best HLA allele-epitope complex, based on docking study (Table 7), is shown in Fig. 1 a–d.

**Table 7** Binding energy determination by autodock

Protein name	Epitope	HLA Class II allele	Binding energy	Intermolecular energy	Internal energy	Torsional energy	Vdw-lbDesolv energy	Electrostatic energy
Nucleocapsid	MRILKTARE	DRB1_0410	4.01	-7.93	-8.58	11.93	-6.28	-1.65
		DRB1_0806	2.66	-9.28	-7.49	11.93	-6.95	-2.33
		DRB1_0405	2.28	-9.66	-7.2	11.93	-9.5	-0.16
	IRIFVPATN	DRB1_1304	3.91	-8.02	-7.12	11.93	-7.44	-0.58
		DRB1_1304	-2.56	-11.51	-7.83	8.95	-11.25	-0.26
		DRB1_0806	-1.88	-10.82	-6.55	8.95	-10.92	-0.1
Matrix	LNLRSRLAA	DRB1_0817	0.13	-10.01	-6.01	10.14	-9.43	-0.59
		DRB1_0701	-2.76	-13.8	-6.82	11.04	-13.86	-0.06
	VRRAGKYYS	DRB1_1301	1.12	-1.32	-7.87	10.44	-9.030	-0.29
		DRB1_0301	-1.01	-11.45	-7.88	10.44	-11.38	-0.07
		DRB1_0421	-0.15	-10.59	-6.86	10.44	-10.02	-0.57
		DRB1_1104	-0.62	-11.06	-5.7	10.44	-9.79	-1.28
Fusion	MILYVLSI	DRB1_1104	1.06	-9.08	-10.05	10.14	-8.78	-0.3
		DRB1_0402	-0.24	-10.38	-8.75	10.14	-10.31	-0.07
	VRLKCLLCG	DRB1_0806	-1.51	-11.66	-8.25	10.14	-11.39	-0.26
		DRB1_0701	-0.73	-11.17	-7.76	10.44	-11.31	-0.14
Glycoprotein	IRLAVRPKS	DRB1_0301	-0.22	-10.37	-6.58	10.14	-10.05	-0.32
		DRB1_1301	3.76	-6.98	-10.91	10.74	-6.68	-0.3
		DRB1_1102	1.53	-10.1	-8.87	11.63	-9.72	-0.38
W protein	LRLREPPQS	DRB1_0806	-0.38	10.53	-7.13	10.14	-10.64	-0.11
V protein	LRLREPPQS	DRB1_0804	-0.76	-10.9	-9.86	10.14	-11.36	-0.46
		DRB1_0806	0.94	-9.2	-7.63	10.14	-9.33	-0.13
		DRB1_0806	2.87	-9.36	-8.99	12.23	-8.98	-0.38
		DRB1_0817	4.28	-7.98	-8.49	12.23	-8.04	-0.06
		DRB1_0804	1.98	-10.25	-7.44	12.23	-10.35	-0.1
Polymerase	FLILNRLLS	DRB1_1305	0.29	-10.45	-6.52	10.74	-10.02	-0.43
		DRB1_1128	-0.55	-11.29	-7.46	10.74	-4.62	-0.16
		DRB1_1101	0.9	-9.84	-8.61	10.74	-9.77	-0.07
	MYMKDKALS	DRB1_0804	2.38	-8.96	-7.89	11.34	-6.71	-2.25
		DRB1_0806	2.13	-9.2	-9.37	11.34	-7.03	-2.18

### Binding Energy Determination of Epitope & HLA Class II Allele

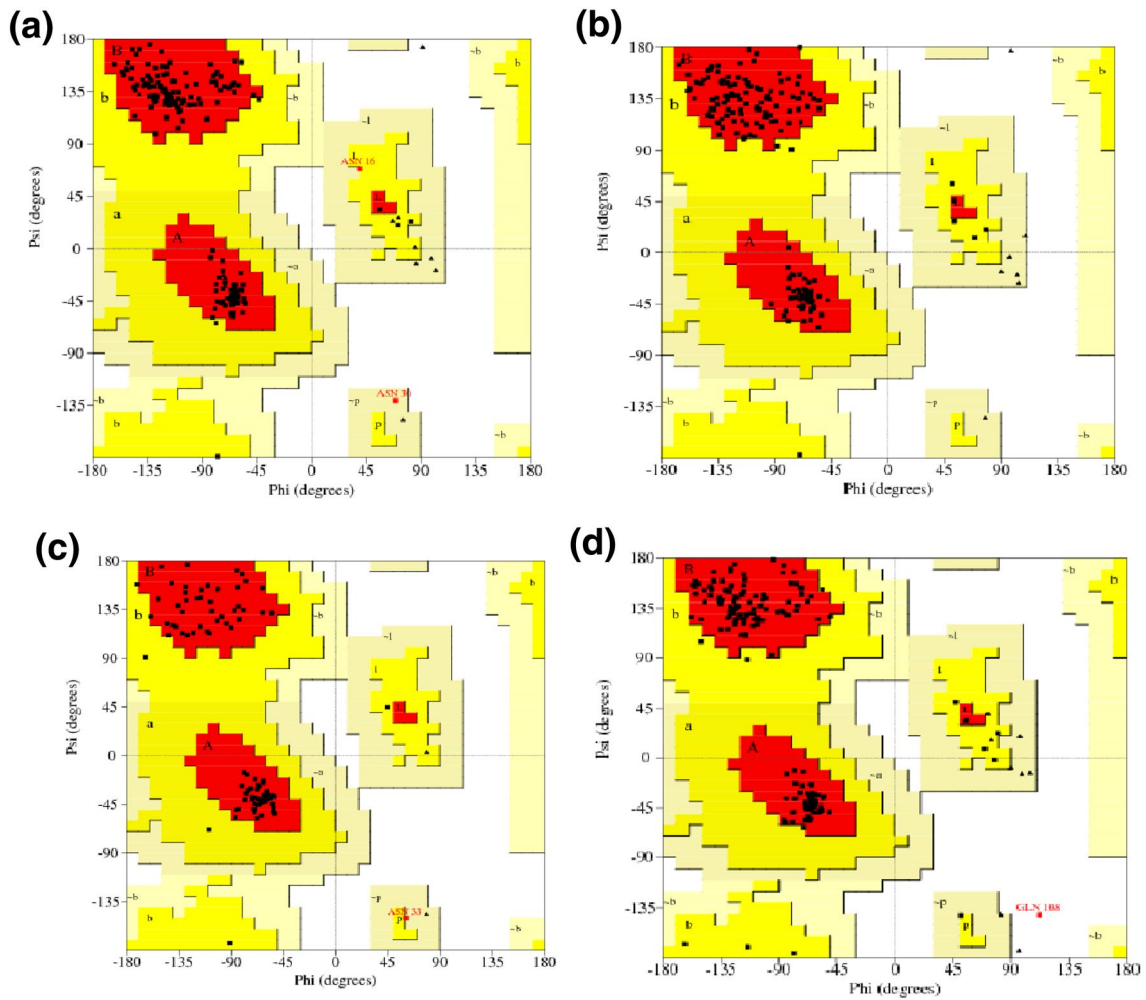
The interaction studies of selected T-cell epitopes with their respective highest ProPred scorer HLA class II allele binders were predicted by utilizing Autodock 4.2 (Table 7). Amongst the docked complexes, nucleocapsid protein peptide IRIFVPATN were dock with two alleles DRB1\*0806 & DRB1\*1304 and nucleocapsid protein peptide MRNLLSQSL with DRB1\*0701 allele formed stable complexes having binding energy value of -1.88 kcal/mol, -2.56 kcal/mol and -2.76 kcal/mol respectively. Matrix protein peptide VRRAGKYYS was docked with DRB1\*0301 HLA allele, it showed binding energy value of -1.01 kcal/mol. Fusion protein peptide VRLKCLLCG

was docked with DRB1\*0806 HLA allele and has shown binding energy value of -1.51 kcal/mol Figs. 2, 3, 4, 5 and 6 respectively. The docked complexes were visualized by Python molecular viewer.

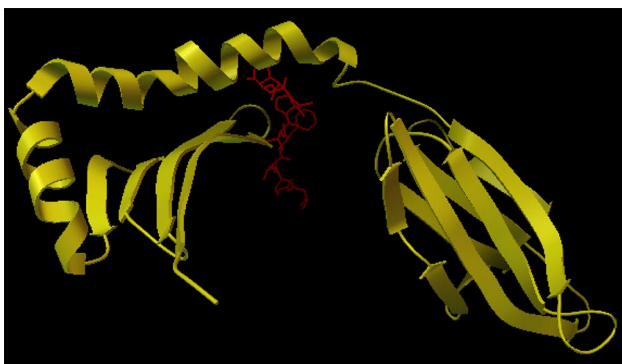
### Docked Complex Stability Validation & RMSD Plot

The lowest binding energy docked complexes of T-cell epitope & HLA Class II allele were subjected to MD simulation by NAMD. The RMSD plot of docked complexes IRIFVPATN-DRB1\*0806, IRIFVPATN-DRB1\*1304, MRNLLSQSL-DRB1\*0701, VRRAGKYYS-DRB1\*0301 and VRLKCLLCG-DRB1\*0806 showed the highest peak at 0.95 Å, 0.98 Å, 0.99 Å, 1.02 Å and 1.01 Å RMSD (Fig. 7 a-c, 8, 9) respectively.

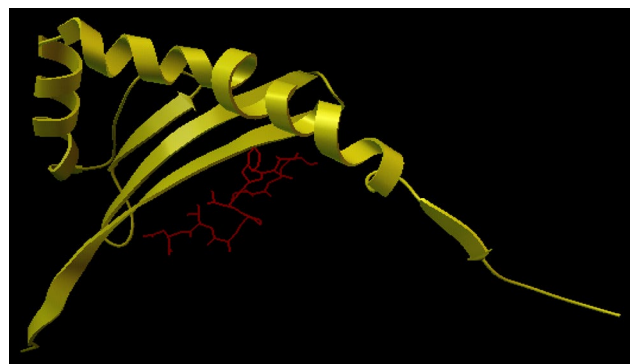




**Fig. 1** Ramachandran plot of protein model: **a** DRB1\*0701, **b** DRB1\*0301, **c** DRB1\*1304, **d** DRB1\*0806

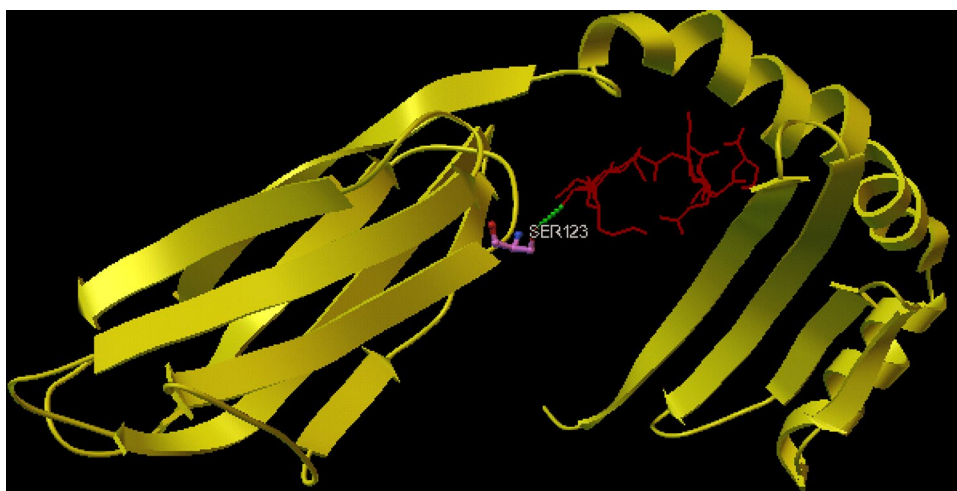


**Fig. 2** Docked complex of Nucleocapsid protein peptide IRIFVPATN with DRB1\*0806 allele

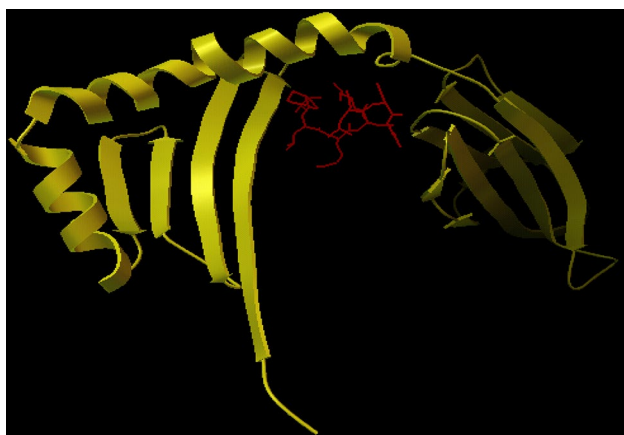
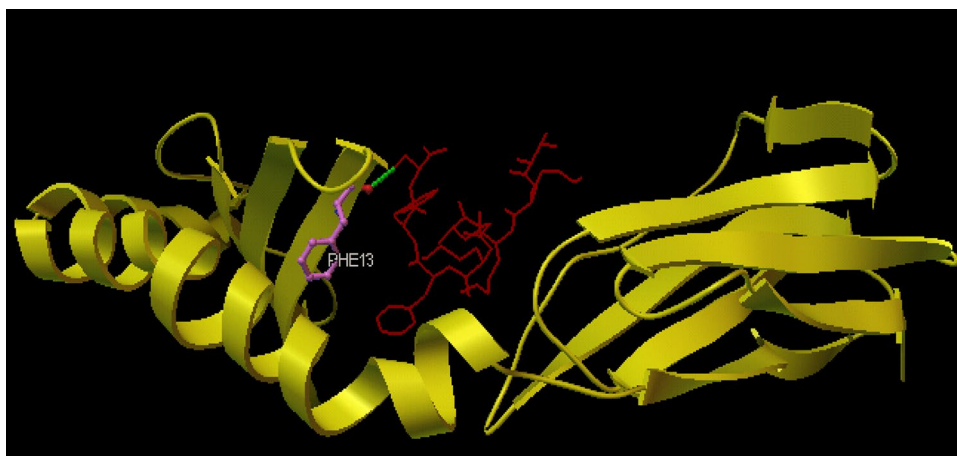


**Fig. 3** Docked complex of Nucleocapsid protein peptide IRIFVPATN with DRB1\*1304 allele

**Fig. 4** Docked complex of Nucleocapsid protein peptide MRNLLS<sub>QSL</sub> and DRB1\*0701 allele. Formation of one H-bond with SER123 (OG) and peptide MET1(N)



**Fig. 5** Docked complex of Matrix protein peptide VRRAGKYYS and DRB1\*0301 allele. Formation of one H-bond with PHE13 (O) and peptide SER9 (N)

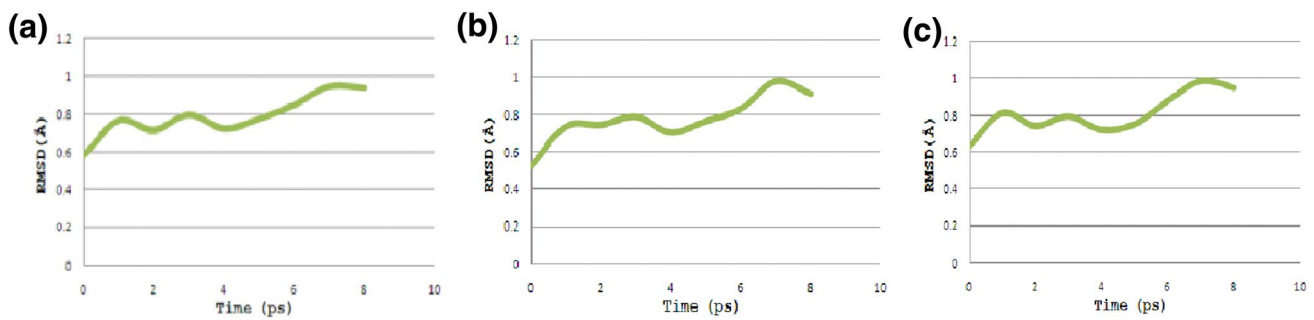


**Fig. 6** Docked complex of Fusion protein peptide VRLKCLLCG with DRB1\*0806 allele

### Population Coverage Estimation of Predicted T-Cell Epitopes

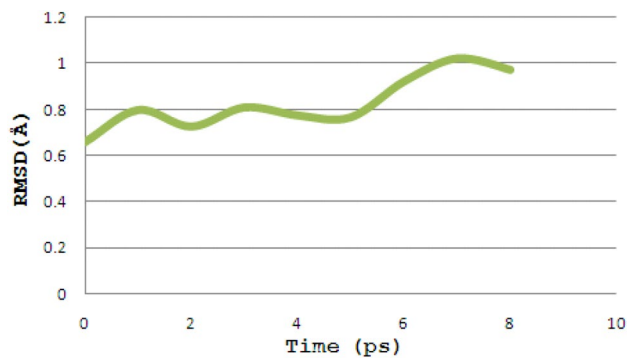
MHCPred (Guan et al. 2003) was employed to predict the potential HLA allele binders for selected T cell epitopes IRIFVPATN, MRNLLS<sub>QSL</sub>, VRRAGKYYS and VRLKCLLCG, along with their interacting affinity with IC<sub>50</sub> value to be  $\leq 500$  nM (Table 2). IEDB tool predicted the four epitopes IRIFVPATN, MRNLLS<sub>QSL</sub>, VRRAGKYYS and VRLKCLLCG have population coverage of 40.24%, 56.78%, 36.82% and 57.08% respectively for total world population (Table 8). Highest coverage of 59.74%, 53.24% and 66.78% for epitope IRIFVPATN, VRRAGKYYS and VRLKCLLCG respectively were found for population of the Northeast Asia. Epitope MRNLLS<sub>QSL</sub> has highest coverage of 70.28% in the population of South Asia. The results of high population coverage suggested that the putative T-cell epitopes cover vast majority of geographic population.



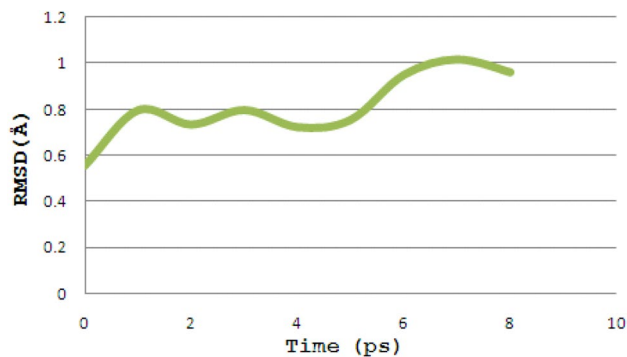


**Fig. 7** Graph displaying MD simulation of Nucleocapsid peptide **a** IRIFVPATN with DRB1\*0806 complex with RMSD highest peak at 0.95 Å. **b** IRIFVPATN with DRB1\*1304 allele complex with RMSD

highest peak at 0.98 Å. **c** MRNLLSQSL and DRB1\*0701 allele complex with RMSD highest peak at 0.99 Å



**Fig. 8** Graph displaying MD simulation of Matrix peptide VRRAG-KYYs and DRB1\*0301 allele complex with RMSD highest peak at 1.02 Å



**Fig. 9** Graph displaying MD simulation of Fusion peptide VRLK-CLLCG with DRB1\*0806 allele complex with RMSD highest peak at 1.01 Å

vaccine candidate against different strains of the infecting organism. MSA result of best predicted epitopes IRIFVPATN and MRNLLSQSL from nucleocapsid, VRRAG-KYYs from Matrix and VRLKCLLCG from Fusion proteins showed that epitopes were well conserved in their entire source protein sequences from different strains of HeV available at NCBI databank. The MSA results were visualized using Jalview (Fig. 10 a–d). Results were further verified by IEDB epitope conservation analysis tool. It was found that predicted epitopes amino acid sequence were 100% conserved amongst all the NCBI protein sequence of the source protein of HeV (Table 3).

## Conclusion

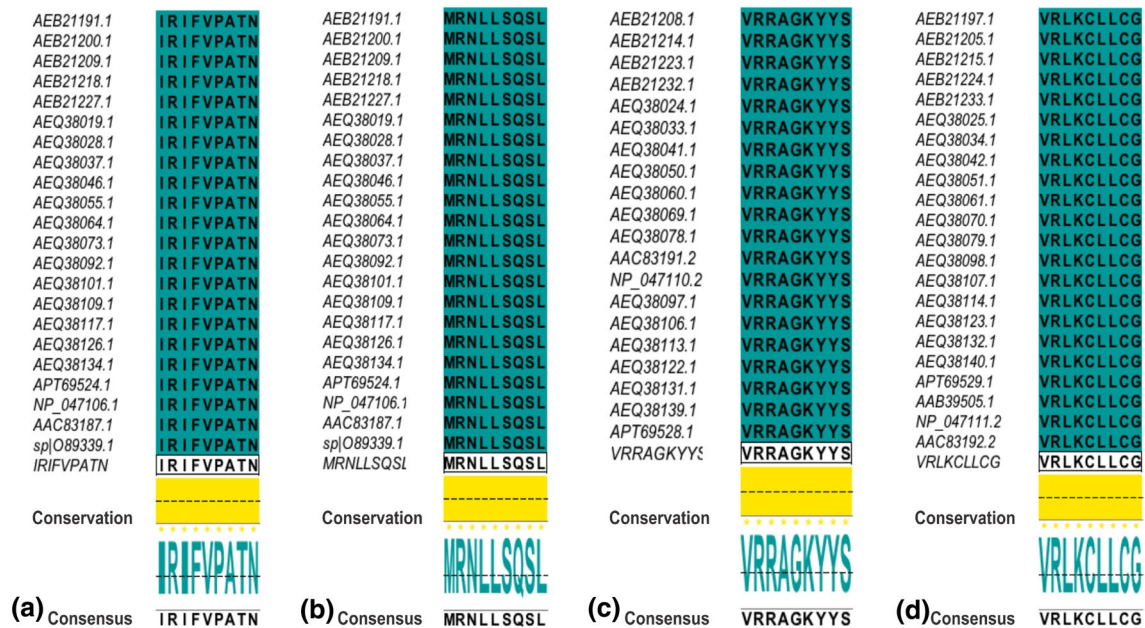
This study identified the potential nanomer T-cell epitopes as a vaccine candidate for Hendra virus. T-cell epitopes IRIFVPATN (nucleocapsid), MRNLLSQSL (nucleocapsid), VRRAGKYYs (matrix) and VRLKCLLCG (fusion) were found to have considerable binding with DRB1\*0806, DRB1\*1304, DRB1\*0701, DRB1\*0301 and DRB1\*0806 HLA class II alleles respectively. The chosen epitopes have shown 100% conservancy throughout all the respective protein sequences from different strains of HeV. These epitopes have shown to have high binding affinity with HLA class II alleles, stable complex formation tendency with HLA class II allele as confirmed by MD simulation results and significant worldwide population coverage. All these results make these epitopes to be a potential candidate for epitope-based vaccine development against HeV infection. Hence reported epitopes may undergo further in-vivo trials to develop vaccine against HeV infection.

## Epitope Conservation and Variability Analysis

The degree of conservation of epitope in protein sequences provides an important insight about its conservancy through evolution and hence its applicability as an epitope-based

**Table 8** Estimated Population coverage of predicted T cell epitopes IRIFVPATN, MRNLLSQSL, VRRAGKYYS and VRLKCLLCG based on MHC-I and MHC-II data using IEDB

Area/population	Epitope (%) for class I and II coverage			
	IRIFVPATN	MRNLLSQSL	VRRAGKYYS	VRLKCLLCG
World	40.24%	56.78%	36.82%	57.08%
East Asia	32.96%	55.25%	34.61%	49.84%
Northeast Asia	59.74%	66.58%	53.24%	66.78%
South Asia	50.55%	70.28%	49.78%	48.22%
Southeast Asia	50.61%	56.60%	46.23%	60.26%
Southwest Asia	29.87%	44.08%	29.28%	43.16%
Europe	41.31%	61.57%	35.32%	59.62%
East Africa	26.00%	25.20%	23.80%	24.39%
West Africa	45.81%	42.46%	28.78%	31.79%
Central Africa	25.19%	36.65%	26.61%	18.14%
North Africa	29.96%	48.82%	34.06%	36.68%
South Africa	39.42%	33.74%	42.58%	21.14%
West Indies	35.01%	49.11%	19.93%	42.58%
North America	35.23%	56.04%	37.08%	57.29%
Central America	0.77%	11.22%	3.65%	0.00%
South America	29.76%	46.12%	27.22%	37.50%
Oceania	40.20%	44.67%	37.35%	47.43%
Average (standard deviation)	36.04% (12.70%)	47.36% (14.46%)	33.31% (11.48%)	41.29% (17.27%)

**Fig. 10** Degree of Conservancy using MSA showed that Epitopes **a** IRIFVPATN (Nucleocapsid) **b** MRNLLSQSL (Nucleocapsid) **c** VRRAGKYYS (Matrix) **d** VRLKCLLCG (Fusion) were 100% con-

served in all the protein sequences obtain from different isolated HeV strains worldwide. Yellow colored rectangle shows conserved nanomer epitopes

**Acknowledgements** The authors are grateful for the necessary immunoinformatic facilities and support provided by the faculty members of Dept. of Zoology, Government P. G. College, Guna & Dept of Microbiology, College of Life Science, Cancer Hospital Campus Gwalior, Madhya Pradesh, India.

**Funding** This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

## Compliance with Ethical Standards

**Conflict of interest** Authors declares that they have no conflict of interest.

**Ethical Approval** This article does not contain any studies with human participants or animals performed by any of the authors.

## References

- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215(3):403–410
- 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(1):153
- Bui HH, Sidney J, Li W, Fusseder N, Sette A (2007) Development of an epitope conservancy analysis tool to facilitate the design of epitope-based diagnostics and vaccines. *BMC Bioinform* 8(1):361
- Doytchinova IA, Flower DR (2007) VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinform* 8(1):4
- Gonzalez-Galarza FF, Christmas S, Middleton D, Jones AR (2010) Allele frequency net: a database and online repository for immune gene frequencies in worldwide populations. *Nucleic Acids Res* 39(suppl\_1):D913–D919
- Guan P, Doytchinova IA, Zygouri C, Flower DR (2003) MHCpred: a server for quantitative prediction of peptide–MHC binding. *Nucleic Acids Res* 31(13):3621–3624
- Gupta S, Kapoor P, Chaudhary K, Gautam A, Kumar R, Raghava GP, Open Source Drug Discovery Consortium (2013) In silico approach for predicting toxicity of peptides and proteins. *PLoS ONE* 8(9):e73957
- Halpin K, Young PL, Field HE, Mackenzie JS (2000) Isolation of Hendra virus from pteropid bats: a natural reservoir of Hendra virus. *J Gen Virol* 81(8):1927–1932
- Hanna JN, McBride WJ, Brookes DL, Shield J, Taylor CT, Smith IL, Craig SB, Smith GA (2006) Hendra virus infection in a veterinarian. *Med J Aust* 185:562–564
- Humphrey W, Dalce A, Schulten K (1996) VMD: visual molecular dynamics. *J Mol Graph* 14(1):33–38
- Jones DT (1999) Protein secondary structure prediction based on position-specific scoring matrices. *J Mol Biol* 292(2):195–202
- Kamthania M, Sharma DK (2016) Epitope-Based Peptides Prediction from Proteome of Nipah Virus. *Int J Pept Res Ther* 22(4):465–470
- Kaur H, Raghava GPS (2004) A neural network method for prediction of  $\beta$ -turn types in proteins using evolutionary information. *Bioinformatics* 20(16):2751–2758
- Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Crystallogr* 26(2):283–291
- Li W, Joshi MD, Singhanian S, Ramsey KH, Murthy AK (2014) Peptide vaccine: progress and challenges. *Vaccines* 2(3):515–536
- Mire CE, Satterfield BA, Geisbert TW (2015) Pathogenesis of Hendra Virus in Humans. In: Singh SK (ed) *Human Emerging and Re-emerging Infections: viral and parasitic infections*, Volume II pp 207–226
- Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, Olson AJ (1998) Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J Comput Chem* 19(14):1639–1662
- Murray K, Rogers R, Selvey L, Selleck P, Hyatt A, Gould A, Gleeson L, Hooper P, Westbury H (1995) A novel morbillivirus pneumonia of horses and its transmission to humans. *Emerg Infect Dis* 1(1):31
- O’Sullivan JD, Allworth AM, Paterson DL, Snow TM, Boots R, Gleeson LJ, Gould AR, Hyatt AD, Bradfield J (1997) Fatal encephalitis due to novel paramyxovirus transmitted from horses. *The Lancet* 349(9045):93–95
- Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skel RD, Kale L, Schulten K (2005) Scalable molecular dynamics with NAMD. *J Comput Chem* 26(16):1781–1802
- Playford EG, McCall B, Smith G, Slinko V, Allen G, Smith I, Field H (2010) Human Hendra virus encephalitis associated with equine outbreak. *Australia 2008 Emerg Infect Dis* 16(2):219
- Robinson J, Halliwell JA, McWilliam H, Lopez R, Parham P, Marsh SG (2012) The imgt/hla database. *Nucleic Acids Res* 41(D1):D1222–D1227
- Šali A, Potterton L, Yuan F, van Vlijmen H, Karplus M (1995) Evaluation of comparative protein modeling by MODELLER. *Proteins: Structure, Function, and Bioinformatics* 23(3):318–3265
- Sanner MF (1999) Python: a programming language for software integration and development. *J Mol Graph Model* 17(1):57–61
- Selvey LA, Wells RM, McCormack JG, Ansford AJ, Murray K, Rogers RJ, Lavercombe PS, Selleck P, Sheridan JW (1995) Infection of humans and horses by a newly described morbillivirus. *Med J Aust* 162(12):642–644
- Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus k, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD, Higgins DG (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using clustal omega. *Mol Syst Biol* 7:539
- Singh H, Raghava GPS (2001) ProPred: prediction of HLA-DR binding sites. *Bioinformatics* 17(12):1236–1237
- Singh S, Singh H, Tuknait A, Chaudhary K, Singh B, Kumaran S, Raghava GP (2015) PEPstrMOD: structure prediction of peptides containing natural, non-natural and modified residues. *Biol Direct* 10(1):73
- Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ (2009) Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25(9):1189–1191

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