

## Immuno-Informatic Speculation and Computational Modeling of Novel MHC-II Human Leukocyte Antigenic Alleles to Elicit Vaccine for Ebola Virus

Arpit Saraswat, Shradha, Amisha Jain, Aakanksha Pathak Sitansu Kumar Verma and Ajay Kumar\*

Department of Biotechnology, Institute of Biomedical Education and Research, Mangalayatan University, Aligarh-202145, India

### Abstract

Ebola hemorrhagic fever (Ebola HF) is a severe, often-fatal and one of the most virulent disease in primates. However the mechanism of escape of virus from the T-cell mediated immune response of the host cell is not explained in any studies yet. In our studies we had aimed on the mapping of novel antigenic determinants of this virus, for impaling the futuristic approach of developing preventive measures against this disease, further we can also study its presumed viral- host mechanism. The complexity of virus- lead compound can be easily accessed through our research work as hereby we emphasized to calibrate the binding energies and stability patterns of antigenic peptides. Our postulation is based on the motifs of MHC II binding epitopes that are specific alleles. Amongst the predicted epitopes, three epitopes were selected including, IVRQRVIPV, FLLMLCLHH and FRLMRTNFL. These three candidates out of 51 antigenic epitopes have the highest score for reactivating with MHC class II in ProPred software and on the basis of highest score binders we are easily be able to distinguish it's binding as well as non-binding peptides, whereas this information is highly significant for designing multi-epitope vaccines (multivalent vaccines) without compromising the human population coverage.

**Keywords:** Ebola HF; Epitope prediction; Vaccine designing; Molecular docking; Multivalent

### Introduction

Ebola virus is one of the two members of a family of RNA viruses called the Filoviridae and of the order Mononegavirales [1]. Ebola HF is an important emerging infection in central Africa and has received much attention in recent years owing to the documented high case-fatality rates (50% to 90%) associated with past outbreaks [2]. Ebola virus was first identified in 1976 when two outbreaks of Ebola hemorrhagic fever (Ebola HF) occurred in northern Zaire (now the Democratic Republic of Congo) and southern Sudan. Out of five identified subtypes of Ebola virus, four of the five have caused disease in humans: Ebola-Zaire, Ebola-Sudan, Ebola-Ivory Coast and Ebola-Bundibugyo. The fifth, Ebola-Reston, has caused disease in nonhuman primates, but not in humans. The most highly virulent subtype of Ebola is Ebola Zaire whose mortality rate is 88%. About 1850 cases with over 1200 deaths have been documented since the Ebola virus was discovered and average fatality rate is 65% [3]. Due to its highly pathogenic nature, scientific research conducted on Ebola must be conducted in a Biosafety Level 4 Lab (BSL-4).

Like all filoviruses, Ebola virions are 80 nm in diameter, but their length varies greatly up to 14,000 nm. It consists of an envelope, a nucleocapsid, a polymerase complex, and a matrix protein. The virus capsid is enveloped, the surface proteins are embedded in a lipid bilayer with surface glycoproteins and the Capsid/nucleocapsid is elongated with helical symmetry. The genome is nonsegmented and contains a single molecule of linear negative-sense, single-stranded RNA [4]. The genome is nearly ~19,000 nucleotides long and bears linearly arranged genes that encode seven structural proteins and one non structural protein. Just as with other viruses there are lipids present and located in the envelope [5]. These filamentous particles may appear in the shape of a shepherd crook or in the shape of a "U" or a "6" and they may be coiled, toroid or branched [6]. The 3' terminus is not polyadenylated and the 5' end is not capped. It codes for 7 structural proteins and one non structural protein. The gene order is 3'-leader -NP-VP35-VP40-GP-VP30-VP24-L-trailer-5' [7].

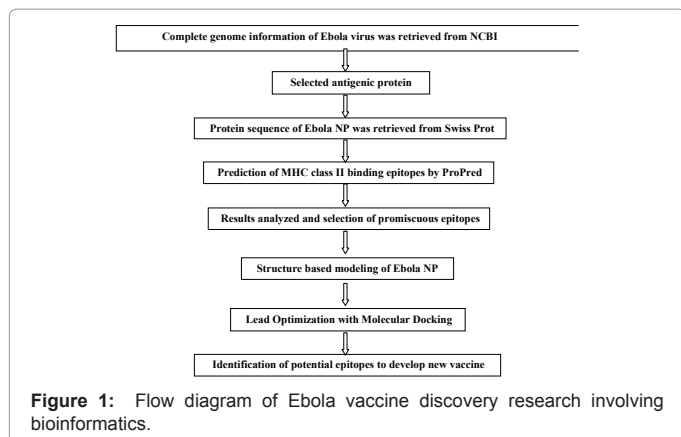
Nucleoprotein of Ebola virus is the largest (79 amino acid residues) nucleoprotein of RNA viruses and can be divided into a hydrophobic N-terminal half (approximately 350 amino acids) and a hydrophilic C-terminal half. Huang et al. showed that Ebola virus NP is O glycosylated and sialylated and that these modifications are required for its interaction with VP35, which suggests that the modifications are important for viral genome replication. NP, together with minor matrix protein VP 24 and polymerase cofactor VP 35, is necessary and sufficient for the formation of nucleocapsid like structures. In recent studies the primary emphasis is on the development of advance yet conventional Ebola vaccine which showed maximum efficacy and stability while shows no or very less side effects Compared to the conventional vaccines, peptide or epitope based vaccines are easy to produce, more specific, and also safe. It is well established at T-cells which plays a critical role in inducing immune response against foreign antigen but for their accuracy an antigenic fragment must bind to MHC molecule. In our research work we are emphasizing on preventing us from this highly infectious disease and thus predicting vaccines and its possible outcomes to model the target proteins and their interrelationships for developing immunity against this virus. This research has enormous public health implications, not only because it might be used to limit this virus, but also because this vaccine strategy may be applied to other highly lethal viruses. The candidate Ebola vaccine is synthesized using modified, inactivated genes from the virus. This gives the immune

\*Corresponding author: Dr. Ajay Kumar, Department of Biotechnology, Institute of Biomedical Education and Research, Mangalayatan University, Aligarh-202145, India, Tel: +91-9412883081; E-mail: [ajaymtech@gmail.com](mailto:ajaymtech@gmail.com)

Received July 06, 2012; Accepted August 13, 2012; Published August 18, 2012

**Citation:** Saraswat A, Shradha, Jain A, Kumar Verma APS, Kumar A (2012) Immuno-Informatic Speculation and Computational Modeling of Novel MHC-II Human Leukocyte Antigenic Alleles to Elicit Vaccine for Ebola Virus. J Vaccines Vaccin 3:141. doi:10.4172/2157-7560.1000141

**Copyright:** © 2012 Saraswat A, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.



system information about viral structures so that it can mount a rapid defense against the real virus ever be encountered.

## Methodology

At the advent of estimating the allocation of the vaccine prediction we had briefly mentioned all the protocol followed by us in the pipeline structure shown in (figure 1).

## Data collection

Nucleoprotein of Ebola virus has been used for immunoinformatic analysis. Complete sequence of Ebola virus NP were acquired from NCBI; accession code, NP\_066243, 739 amino acids.

## Prediction of antigenic T-cell epitopes

Several T-cell epitopes that are identified by more than one MHC and recognized by more than one T-cell clone are called “promiscuous” epitopes. Different prediction methods according to biological information are considered for epitopes. In this study, we have chosen ProPred software for prediction of promiscuous T-cell epitopes. This immunoinformatic tool allows users to predict MHC class II binding regions in antigenic sequence. 51 MHC class II epitope mapping matrices were developed therefore it is helpful for the user in locating promiscuous or allele specific binding regions in a query antigen. The peptides (epitopes) predicted in this software are nanomers. In each nanomer there is an anchor or starting residue that helps to identify those antigenic determinants peptides in NP which binds to several MHC class II molecules with good binding affinity. All these peptides were predicted on the 3% threshold value with highest binding score with HLA alleles. The algorithms predict binding of nanomeric peptides to HLA alleles. Those nanomeric peptides have highest binding affinity and maximum coverage of HLA alleles was selected.

## Structure-based modeling of T-cell epitopes

Molecular modeling and structural analysis [8] was performed for the detection of peptide binding to their respective class II HLA alleles. Sample peptides of high affinity binders for a few alleles where structures are known (DRB1\_1301, DRB1\_0405, and DRB1\_1321) were modeled employing their respective structural templates (1A6A, 1BX2, 1SYS). We have selected three peptides viz IVRQRVIPV, FRLMRTNFL and FLLMLCLHH with the help of scoring based algorithms of ProPred. These peptides have higher binding affinity which was estimated by analyzing the intra-molecular hydrogen bonds, electrostatic, Van der Walls and hydrophobic interactions with the protein residues in the vicinity or closed proximity.

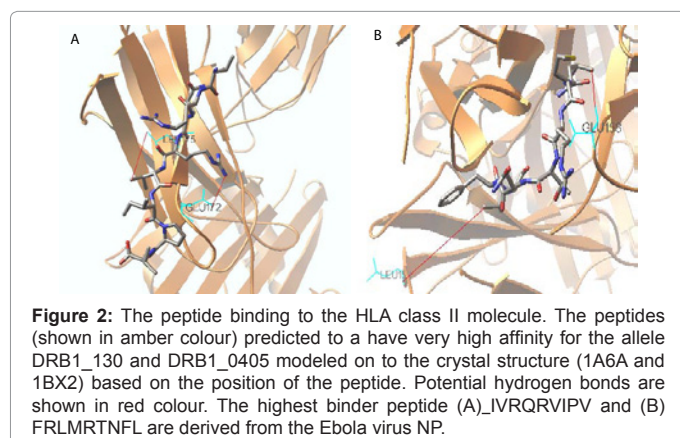
The modeler [9] was used for designing the structures of those alleles whose structures were not available in the PDB server while the CPH model server [10] was used to design the structures of the predicted binding peptides. After designing the structures we performed docking of selected alleles and peptides with the help of Autodock for finding out the energy minimization [11]. Then PMV (Python Molecular

Allele	1	2	3	4	5
DRB1_0101	28.33	33.33	56.67	11.17	0
DRB1_0102	45.00	33.33	56.67	27.83	05.00
DRB1_0301	49.47	-	30.21	42.11	30.53
DRB1_0305	29.67	25.05	31.98	16.48	13.19
DRB1_0306	38.41	20.23	29.55	25.00	25.00
DRB1_0307	38.41	20.23	29.55	25.00	25.00
DRB1_0308	38.41	20.23	29.55	25.00	25.00
DRB1_0309	38.95	17.89	40.74	31.58	20.00
DRB1_0311	38.41	20.23	29.55	25.00	25.00
DRB1_0401	33.49	42.79	65.12	17.44	11.63
DRB1_0402	32.29	27.92	44.79	27.08	21.88
DRB1_0404	35.00	36.14	63.64	03.41	0
DRB1_0405	25.32	57.45	<b>80.85</b>	04.26	0
DRB1_0408	23.64	47.50	75.00	0	0
DRB1_0410	35.96	46.81	70.21	14.89	08.51
DRB1_0421	43.91	34.44	72.89	33.33	18.89
DRB1_0423	35.00	36.14	63.64	03.41	0
DRB1_0426	33.49	42.79	65.12	17.44	11.63
DRB1_0701	41.38	31.03	64.83	41.38	15.52
DRB1_0703	41.38	31.03	64.83	41.38	15.52
DRB1_0801	09.30	51.16	47.67	02.33	12.79
DRB1_0802	06.25	39.75	38.75	0	02.50
DRB1_0804	18.75	27.25	26.25	01.25	15.00
DRB1_0806	20.93	39.53	36.05	13.95	24.42
DRB1_0813	15.86	43.45	58.62	01.15	02.30
DRB1_0817	23.76	61.39	43.56	17.82	10.89
DRB1_1101	32.53	55.18	36.14	13.25	07.23
DRB1_1102	40.48	27.14	34.52	34.52	44.05
DRB1_1104	44.58	43.13	24.10	25.30	19.28
DRB1_1106	44.58	43.13	24.10	25.30	19.28
DRB1_1107	40.66	14.07	20.99	27.47	24.18
DRB1_1114	28.57	39.05	46.43	22.62	32.14
DRB1_1120	38.64	30.68	55.23	38.64	38.64
DRB1_1121	40.48	27.14	34.52	34.52	44.05
DRB1_1128	42.53	45.98	45.52	29.89	14.94
DRB1_1301	<b>50.00</b>	19.32	43.86	50.00	50.00
DRB1_1302	38.64	30.68	55.23	38.64	38.64
DRB1_1304	41.11	38.89	43.33	44.44	51.11
DRB1_1305	42.53	45.98	45.52	29.89	14.94
DRB1_1307	16.18	40.88	39.71	0	08.82
DRB1_1311	44.58	43.13	24.10	25.30	19.28
DRB1_1321	33.71	<b>65.17</b>	44.94	24.72	16.85
DRB1_1322	40.48	27.14	34.52	34.52	44.05
DRB1_1323	28.57	39.05	46.43	22.62	32.14
DRB1_1327	50.00	19.32	43.86	50.00	50.00
DRB1_1328	50.00	19.32	43.86	50.00	50.00
DRB1_1501	35.71	28.57	44.49	60.00	43.88
DRB1_1502	25.51	38.78	54.69	49.80	33.67
DRB1_1506	35.71	28.57	44.49	60.00	43.88
DRB5_0101	08.16	54.08	60.20	07.14	-
DRB5_0105	08.16	54.08	60.20	07.14	-

**Table 1:** Comparison scores between the five epitopes according to prored software.

Peptide	Energy (kcal/mol)	Hydrogen bonds	Site of amino acid	Score	Alleles
IVRQRVIPV	-31.45	2	34-42	50.00	DRB1_1301
FLLMLCLHH	-26.85	2	72-80	80.85	DRB1_0405
FRLMRTNFL	-29.04	2	200-208	65.17	DRB1_1321

**Table 2:** The conformational properties of the peptides with efficient binding energy and present on the variable regions of predicted peptide as investigated by molecular dynamics simulation using Autodock tool v3.0.



Viewer) was used for the visualization of binding position of hydrogen bonding between the selected peptides and alleles.

## Results and Discussion

### Analysis of the MHC class II T cell Epitopes

The promiscuity of binding of a peptide to HLA alleles is important since inclusion of such peptides in the vaccine construct provides a greater population coverage which helps to short out the promiscuous peptide that needs to be in vaccine developments [12].

ProPred web-based MHC class II binding peptide prediction tool was used to narrow down the search for putative and promiscuous antigenic peptides in NP protein of the Ebola virus for the use in development of diagnostic tests and vaccines. CD4+ T-cells play important roles in producing protective immunity [13]. Therefore, prediction antigenic epitopes in the context of MHC class II molecules expressed on APC (antigen presenting cell) recognized by CD4+ T-cells can be used as a suitable tool for vaccination.

NP is the largest, highly conserved Ebola viral protein having major role in capsid assembly and viral replication. It is composed of 739 amino acids having molecular weight of 83.3kd and was analyzed at a 3% threshold value with highest binding score with MHC class II molecules. Analysis of MHC class II binding is based on motifs associated with epitopes or binders for specific alleles. In this study, we selected 3 highest scoring putative epitopes from specific regions. The non-overlapping top 5 nonamers of NP having highest score for reactivating MHC class II has been selected based on their binding affinities with HLA alleles (table 1). Out of these 5, the best 3 epitopes showing maximum binding score with 51 HLA alleles were screened and modeled. The first epitope (IVRQRVIPV) contained 37, the second (FLLMLCLHH) contained 37 and the third (FRLMRTNFL) contained 49 alleles out of the 51 HLA alleles.

### Structure based modeling of HLA-peptide complex

Simulation studies of the epitopes IVRQRVIPV and FRLMRTNFL

formed stable HLA-peptide complexes with the energy minimization of -31.45 (kcal/mol) and -29.04 (kcal/mol) respectively. The other peptide FLLMLCLHH identified in the present study was found antigenically variable with energy minimization value of -26.85 (kcal/mol) and their position and score is shown in (table 2). This can possibly be targeted for designing of vaccine against Ebola virus. The peptides binding to the HLA class II molecule were shown in (figure 2).

## Conclusion

With the advent in the field of computational immunology it is now possible to drastically reduce the time for identification of putative and promiscuous antigenic peptides. The present study was undertaken with an objective to initiate the use of immunoinformatics. These approaches are currently used for prediction of antigenic determinants in the protein sequence of NP of Ebola virus without using their cultures. The prediction of Ebola (NP) virus nanomer epitopes for T cells is recognized against MHC class II molecules. The predicted epitopes may be served as a useful diagnostic reagent for evaluating T-cell responses in the context of natural infection and also might be helpful for designing a subunit vaccine against Ebola virus.

## Acknowledgement

Authors are grateful to Dr. G.K. Aseri (Director, I.B.M.E.R., Mangalayatan University Aligarh, UP, India) for providing necessary facilities and encouragement. The authors are also thankful to all faculty members of the Institute of Biomedical Education and Research, Mangalayatan University Aligarh, UP, India for their generous help and suggestions during the course of experimental work and manuscript preparation.

## References

- Bennett D, Brown D (1995) Ebola virus. *BMJ* 310: 1344-1345.
- Shears P (2000) Emerging and reemerging infections in africa: the need for improved laboratory services and disease surveillance. *Microbes Infect* 2: 489-495.
- World Health Organization (2011) Ebola hemorrhagic fever.
- The Universal Virus Database Taxonomy on viruses. International Committee version 4.
- Feldmann H, Klenk HD, Sanchez A (1993) Molecular biology and evolution of filoviruses. *Arch Virol Suppl* 7: 81-100.
- Kiley MP, Bowen ET, Eddy GA, Isaacson M, Johnson KM, et al. (1982) Filoviridae: a taxonomic home for Marburg and Ebola viruses? *Intervirology* 18: 24-32.
- Taylor DJ, Leach RW, Bruenn J (2010) Filoviruses are ancient and integrated into mammalian genomes. *BMC Evol Biol* 10: 193.
- Chaitra MG, Hariharaputran S, Chandra NR, Shaila MS, Nayak R (2005) Defining putative T cell epitopes from PE and PPE families of proteins of *Mycobacterium tuberculosis* with vaccine potential. *Vaccine* 23: 1265-1272.
- Eswar N, Webb B, Marti-Renom MA, Madhusudhan MS, Eramian D, et al. (2007) Comparative protein structure modeling using MODELLER. *Curr Protoc Protein Sci* 2: Unit 2.9.
- Nielsen M, Lundegaard C, Lund O, Petersen TN (2010) CPHmodels-3.0--remote homology modeling using structure-guided sequence profiles. *Nucleic Acids Res* 38: W576-581.
- Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, et al. (1998) Automated Docking Using a Lamarckian Genetic Algorithm and an Empirical Binding Free Energy Function. *J Computational Chemistry* 19: 1639-1662.
- Herrera DIM, Morales JAM, Cardena AEP, Sanchez BM, Chessani MAR, et al. (2010) Use of RB51 vaccine for small ruminants brucellosis prevention, in Veracruz, Mexico. *Int J Dairy Sci* 5: 10-17.
- Shi Y, Wu C, Zhou WY, Mao XH, Guo G, et al. (2007) Identification of H-2d restricted Th epitopes in Urease B subunit of *Helicobacter pylori*. *Vaccine* 25: 2583-2590.