Designing of an Epitope-Based Universal Peptide Vaccine against Highly Conserved Regions in RNA Dependent RNA Polymerase Protein of Human Marburg Virus: A Computational Assay

Article *in* Anti-Infective Agents · July 2019

DOI: 10.2174/2211352517666190717143949

CITATIONS
3

READS
131

5 authors, including:

S. M. Neaz Mahmud
Mawlana Bhashani Science and Technology University
10 PUBLICATIONS **13** CITATIONS
SEE PROFILE

Md. Mahbubur Rahman
Medinova Medical Services Ltd.
4 PUBLICATIONS **3** CITATIONS
SEE PROFILE

Antora Kar
Mawlana Bhashani Science and Technology University
4 PUBLICATIONS **10** CITATIONS
SEE PROFILE

Md. Arif Khan
University of Development Alternative
73 PUBLICATIONS **398** CITATIONS
SEE PROFILE

Send Orders for Reprints to reprints@benthamscience.net

Anti-Infective Agents, **2019,** *17,* **000-000**

RESEARCH ARTICLE

Designing of an Epitope- Based Universal Peptide Vaccine against Highly Conserved Regions in RNA Dependent RNA Polymerase Protein of Human Marburg Virus: A Computational Assay

S.M. Neaz Mahmud[1,*], Md. Mahbubur Rahman[1], Antora Kar[1], Nasreen Jahan[1] and Md. Arif Khan[2,3,*]

[1]*Department of Biotechnology and Genetic Engineering, MawlanaBhashani Science and Technology University, Santosh, Tangail-1902, Bangladesh;* [2]*Department of Biotechnology and Genetic Engineering, University of Development Alternative, 4/4B, Block A, Lalmatia, Dhaka-1209;* [3]*Bio-Bio-1 Research Foundation, Sangskriti Bikash Kendra Bhaban, 1/E/1, Poribag, Dhaka 1000, Bangladesh*

ARTICLE HISTORY

Received: March 28, 2019
Revised: May 30, 2019
Accepted: June 21, 2019

DOI:
10.2174/2211352517666190717143949

Abstract: ***Introduction*****:** Marburg viruses are a group of negative-stranded RNA virus. It was first identified in 1967 during a small outbreak. During that outbreak, the fatality rate increased highly and so many people died by the Marburg virus. Later seven strains of Marburg virus were identified from those infected humans. This virus causes Marburg Virus Disease (MVD) in human referred to as Marburg hemorrhagic fever. Marburg virus is endemic only to Africa; however, there have been outbreaks in Europe and the U.S.A. in recent times.

Objective**:** However, the Marburg virus has a high fatality rate, so a preventive measure should be taken to prevent infection. As there is no effective therapeutic agent available against these viruses, effective vaccine design touching all strains would be a great step for human health.

Methods**:** In our recent study, we used *in silico* analysis for designing a novel epitope-based vaccine against all strains of Marburg virus. As it consists of several structural proteins and multiple sequence alignment (MSA) of Glycoproteins, RNA-directed RNA polymerases, Nucleoproteins, Vp24 proteins, Vp30, Vp35, and Vp40 proteins showed all strains of Marburg virus were conserved in RNA-directed RNA polymerase proteins. Using that protein's conserved region, T-cell and B-cell epitopes were determined.

Results**:** Among the predicted epitope, only TIGNRAPYI was found to be highly immunogenic with 100% conservancy among all strain of human Marburg virus. The analysis also showed both types I and II major histocompatibility complex molecules interact with this epitope and found to be nonallergenic too.

Conclusion**:** *In vivo* study of the proposed peptide is suggested for novel universal vaccine production that might be an effective way to prevent human Marburg virus disease.

Keywords: Marburg virus, marburg hemorrhagic fever, universal peptide vaccine, RNA directed RNA polymerase, epitope.

1. INTRODUCTION

Marburg viruses are a member of filoviridae family, Marburg virus genus, and Mononegavirales order [1]. Marburg viruses are closely related to Ebola virus (EBOV) that causes serious infection on human and primates. This virus is an enveloped, single-stranded, unsegmented, negative-sense RNA virus which contains a filamentous structure that can appear in a shape like which is a 6 or spiralled like a Snail and can sometimes be branched [2]. The Virions of the infected virus are 80nm in diameter and 800nm in length, although the length can vary up to 1400nm [3] and associated with peak infective is 790nm. They contain long non-coding region at their 3′ or 5′ ends which probably contributes to the stability of viral transcript. Marburg virus consists of several structural proteins and a negative-stranded linear RNA genome, about 19 kb in size [4]. In the center, it consists of a helical ribonucleocapsid, which may contain the genomic RNA wrapped around a polymer of nucleoprotein (NP) [5]. The RNA dependent RNA polymerase (L) protein lies in association with ribonucleoprotein along with the polymerase cofactor (vp35) and a transcription activator (vp30) [6]. The nucleoprotein is embedded with a matrix, formed by the major (vp40) and minor (vp24) matrix proteins [7]. These particles are surrounded by a lipid membrane derived from the host cell membrane. The membrane anchors a glycoprotein (GP1, 2) that projects 7 to 10 nm spikes away from its surface.

*Address correspondence to this author at the Department of Biotechnology and Genetic Engineering, University of Development Alternative, 4/4B, Block A, Lalmatia, Dhaka-1209 and Bio-Bio-1 Research Foundation, Sangskriti Bikash Kendra Bhaban, 1/E/1, Poribag, Dhaka 1000, Bangladesh;
E-mail: arifkhanbge35@gmail.com

2211-3525/19 $58.00+.00

First Marburg virus was identified in 1967 [8]. Some small outbreaks of Marburg virus were noticed in some German cities and during the outbreak, 31 people were infected and 7 people died [9]. Since then, different strains of Marburg virus have emerged and Lake Victoria Marburg virus strain (Popp67, Musoke 80, Ravn-87, Ozolin-75, Angola-2005, Cis-67, and Marburg virus) was identified as infecting humans. Marburg virus causes Marburg virus disease (MVD) in humans referred to as Marburg hemorrhagic fever [10]. After the first outbreak, the virus spread in Europe, Africa, and the U.S.A. [11-14]. In 1987-88, it created serious effects in Kenya, Koltsovo, and the Soviet Union and caused 100% death [15]. During 2004 in Angola, several hundreds of people were infected by the Marburg virus and the fatality rate went up to 90% [16]. It also caused an outbreak in the U.S.A, Netherland, and Uganda in 2008 [17]. From the clinical experience of investigation, it was found that there were no effective therapeutic agent, drugs, and vaccines developed against Marburg virus. Several common therapeutic agents are used in the treatment of Marburg virus. As there was no therapeutic agent available against this virus, the identification of therapeutics has high priority because of high fatality rates. As the virus is now becoming pathogenic and sometimes cause outbreaks so precautionary steps should be taken to prevent human death [18]. On the other hand, the vaccine treatments are now becoming popular and very effective against various infectious diseases and several other diseases. Vaccination is very effective to control mumps, chickenpox, yellow fever, Japanese encephalitis, rabies, rubella, influenza, smallpox, hepatitis A and B, *etc*. [19]. So designing a vaccine against conserved regions of all Marburg viruses is a major challenge at present.

Recently, viral vaccine preparation in epitope base vaccine designing method is the most challenging approach because of the availability of a large number of sequence information [20, 51, 52]. The immunological approaches of bioinformatics are known as immunoinformatics and are spreading very quickly towards the world. Design and identification of T- cell epitope, B- cell epitope and Human Leukocyte Antigen (HLA) ligands are mainly performed in immunoinformatics [21-25, 53, 54]. Moreover, computational epitope prediction and vaccine design may reduce time and cost. In our proposed study, we designed a putative epitope-based vaccine against Marburg virus. Viral proteins are analysed using bioinformatics methods to catch conserve peptide region and mapped an evolutionarily conserved epitope. By using this method the 3D structure of RNA directed RNA polymerase is also determined. So the vaccine we designed may be used as a universal vaccine against all types of Marburg virus.

2. MATERIALS AND METHODS

2.1. Retrieving Marburgvirus Structural and Non-Structural Protein Sequences

A total of 41 polyproteins were used to design this vaccine. We found 7 glycoproteins, 6 RNA-directed RNA polymerases, 6 Nucleoproteins, 6 Vp24 proteins, 5 Vp30, 5 Vp35, and 6 Vp40 protein sequences from UniProt KB [26] protein database.

2.2. Identification of Conserved Sequences

EBL-clustalW [27] program was used to identify Conserved sequences, which aligned the retrieved sequences. Gonnet matrix [27] was used in multiple sequence alignment (MSA). Jalview v2, [28] tool was used to retrieve the alignment. We estimated protein variability index from protein variability server (PVS) [29] applying Wu-Kabat Variability coefficient. Conserved sequences were identified from multiple sequence alignment in which similar amino acid sequences were highest in number and no gap was found towards the sequences. Antigenicity of the conserved sequences was predicted after that.

2.3. Antigenicity Determination of Conserved Sequences

Antigenicity of conserved sequences was identified by VaxiJen v2.0, [30] a Web-based server. In this research, the prediction was done using default parameters, with a threshold value of 0.5.

2.4. T-cell Epitope Prediction

We used NetCT [31] and CTLPred [32] two online server to predict T- cell epitopes. Primarily, the T-cell epitopes were identified in the NetCTL server. This epitope prediction process follows some steps like major histocompatibility complex class I (MHC-I) binding, proteasomal C terminal cleavage and transporter of antigen peptide (TAP) transport efficiency and all these were performed to identify the epitopes [12]. MHC-I supertype was used only for epitope prediction and for Proteasomal cleavage and MHC-I binding procedure artificial neural networks were applied. TAP transport efficiency was determined in the weight matrix. We found 0.89 and 0.94 as the value of sensitivity and specificity of epitope at a 0.5 threshold level. Finally, CTLPred was applied with default parameters for further confirmation about epitopes.

2.5. MHC-I and MHC-II Restriction Analysis

Prediction of MHC-I [33] and MHC-II [34, 35] alleles that bind with proposed epitopes, accomplished by the T Cell Epitope Prediction Tools from Immune Epitope Database and Analysis Resource (IEDB-AR). Half-maximal inhibitory concentration (IC_{50}) of the peptide binding to MHC-I molecules was estimated by Stabilized Matrix Method, [36] employing preselected 9.0-mer epitope. For MHC-II binding analysis, IEDB-recommended method was used for specific HLA-DQ, HLA-DP, and HLA-DR loci. Antigenic conservancy and MHC-I analysis were utilized for choosing specific peptides used for the prediction of MHC-II interaction.

2.6. B-cell Epitope Prediction

By interacting with B lymphocytes B-cell epitope stimulates immune response and then it differentiates B lymphocytes into plasma and memory cells [37]. Bepipred (v1.0) [38] and Immune Epitope Database (IEDB) [39], a B cell epitope prediction tool, were used to predict B-cell epitope. Using a hidden Markov model [38], Bepipred predicts linear B-cell epitopes. Default 0.35 was used as a threshold value for Bepipred analysis. For conformation IEDB (B cell epitope prediction tool) was used for linear B cell epitope prediction. Finally, Kolaskar and Tongaonkar Antigenicity [40] approaches were used at the threshold level of 1.000, to predict the antigenic B-cell epitopes from the provided protein sequence with high accuracy.

2.7. Prediction of Surface Accessible Epitopes

Surface accessibility of predicted epitopes from the conserved sequences was identified using, Emini surface accessibility [41] prediction tool of the B cell epitope prediction tools of The Immune Epitope Database (IEDB) at threshold level 1.0.

2.8. Prediction of Epitope Hydrophilicity

Parker hydrophilicity [42] prediction tool of Immune Epitope Database (IEDB) was used to determine the hydrophilicity of the predicted epitopes from conserved sequences. Default threshold 3.448 was used for this approach.

2.9. Allergenicity and Epitope Conservancy Analysis

Allergenicity of the proposed epitopes was determined using a Web-based server AlgPred [43] with high accuracy. Allergenicety of the epitopes was predicted with an accuracy of about 85% at a threshold value of −0.4, using hybrid prediction (SVMc + IgEepitope + ARPs BLAST + MAST) method. The guidelines of the Food and Agriculture Organization/World Health Organization, 2003 were maintained in this prediction process. Epitope conservancy predictions of the proposed epitopes were done by epitope conservancy [44] analysis tool from IEDB analysis resource. Protein sequences were searched for identities to calculate the conservancy level of the epitopes.

2.10. Homology Modelling and Evaluation of Protein 3D Model

As no experimental structure of RNA directed RNA polymerase protein of any human Marburg virus was found in Protein Data Bank (PDB). So homology model of the 3D structure of the conserved sequence was obtained by I-TASSER [45] server. I-TASSER server uses multiple threading alignments to give protein 3D structure. Then PROCHECK [46] software was used to verify I-TASSER provided top models quality. The best model was selected from PROCHECK analysis by following the criteria that the G factor was highest, and amino acid residues are in the favorable region was higher. Then UCSF Chimera [47] visualization tool was used to detect the epitope in the best model.

3. RESULTS

3.1. Conserved Sequences from the Retrieved Sequences

Conserved sequences were identified through the analyses of MSA by using ClustalW [27] and protein variability index [29]. From the analyses of MSA, it was founded that human Marburg viruses were conserved in the RNA directed RNA polymerase (Additional file 1: Fig. **S1**). MSA of the G, N, Vp24 protein, Vp30, Vp35, and Vp40 proteins showed no significant conserved sequences (Additional file 2: Fig. **S2**, Additional file 3: Fig. **S3**, Additional file 4: Fig. **S4**, Additional file 5: Fig. **S5**, Additional file 6: Fig. **S6**, Additional file 7: Fig. **S7** respectively). Total 309 amino acids of 3 conserved regions among all human Marburg viruses were founded by the MSA (Fig. **1**) and protein variability index (Fig. **2**) from the RNA directed RNA polymerase proteins. The protein variability index of the conserved peptide was determined by using the PVS server (Fig. **2**). The conservancy threshold was 1.0 in this analysis.

3.2. Determination of the Antigenic Peptides

At first, four conserved regions were identified from the RNA directed RNA polymerase-L according to the highest number of identical amino acids in the MSA (MSA number: 537-652 and MSA number: 1207-1584). These conserved sequences were then used to determine their antigenicity by the VaxiJen v2.0 [30] server. From the VaxiJen score, top two conserved peptides were (MSA number: 1207-1334 and 537-652) selected for further analysis (Table **1**). The peptide with best VaxiJen score (MSA: 1207–1334) showed better results during further analysis of T-cell epitope identification, MHC interaction analysis, and B-cell epitope identification was selected for epitope prediction.

3.3. Specification of T-cell epitopes and MHC restriction analysis

Combined approaches of 12 MHC-I supertypes were used to identify T-cell epitopes by NetCTL [31] server. Five epitopes with top scores were selected from the combined results of NetCTL server (Table **2**).

Five T-cell epitopes were also identified by CTLPred [32] server through artificial neural networks and support vector machines combined approach (Table **3**). Both servers predict nine mer epitopes.

Epitopes which were common in both servers were chosen and applied for the MHC-binding analysis. MHC-I alleles which showed higher affinity (IC50 < 200nM) to propose T-cell epitopes through Stabilized Matrix Method, are listed in (Table **4**). MHC-II alleles binding to the proposed T-cell epitopes are also listed in (Table **4**).

3.4. Determination of B-Cell Epitope

Conserved protein sequences were used to predict B-cell epitopes from Immune Epitope Database (IEDB) [39] B-cell epitope prediction tool and Bepipred [38]. Bepipred predicts linear B-cell epitope at threshold 0.35. Nine epitopes were predicted by Bepipred which were most consistent to B-cell epitopes. The results are given in (Table **5**).

Potential B-cell epitopes must be highly antigenic. Antigenicity of the peptides was determined by using Kolaskar and Tongaonkar antigenicity [40] prediction tool in Immune Epitope Database (IEDB) B-cell epitope prediction tool. 12 epitopes were found as antigenic at threshold level 1.00 (Table **6** and Fig. **3**).NRAPYIG epitope was found to be common in both prediction tools.

To induce an immune response B-cell epitope must be surface accessible so that antibody can bind to the epitope. Emini surface accessibility [39] prediction tool from the Immune Epitope Database (IEDB) was used to predict surface accessible B-cell epitopes from the conserved peptide. We found eight surface accessible epitopes above the threshold level (Table **7**) from the analysis. In this analysis threshold was set at 1.00 with maximum propensity score of the peptide was 7.898 and minimum score was 0.112 (Fig **4**).

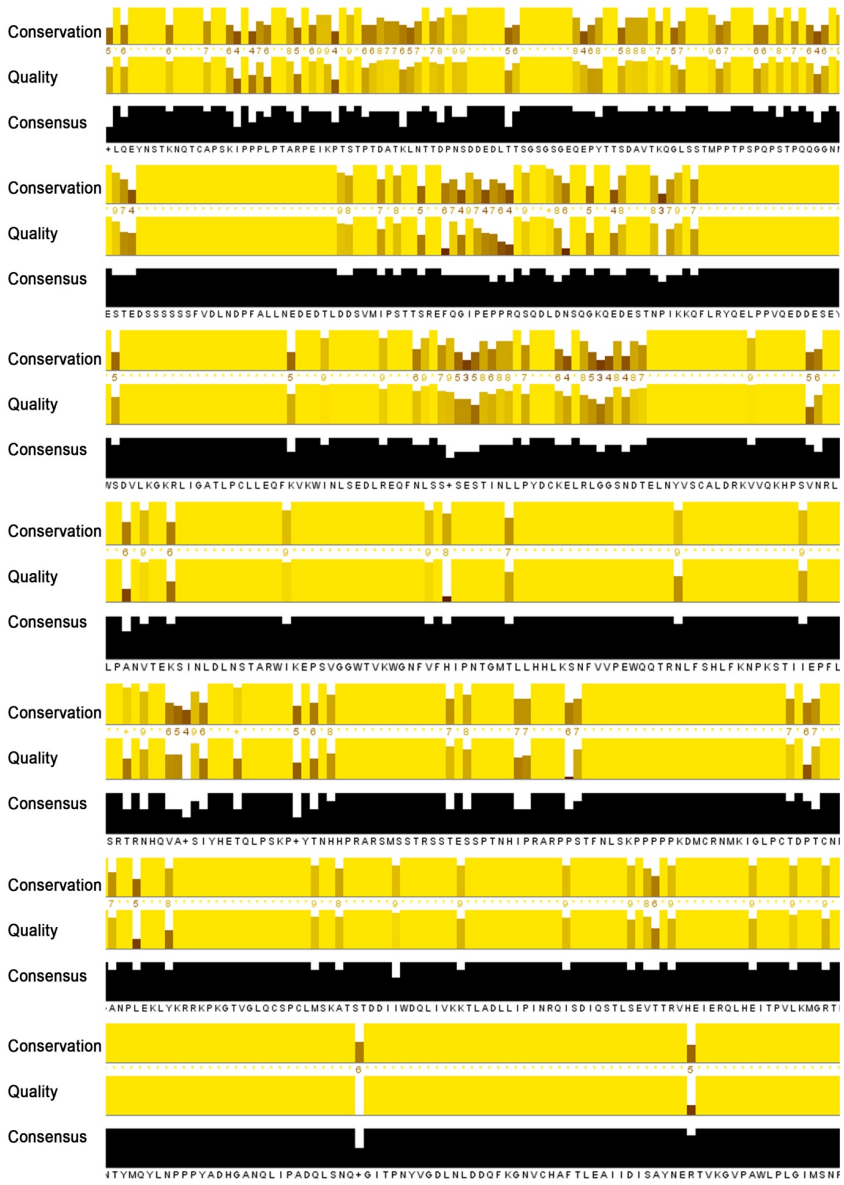

Fig. (1). Most conserved regions of RNA directed RNA polymerase. RNA directed RNA polymerase of human Marburg Virus is most conserved among all proteins and multiple sequence alignment of 6 RNA directed RNA polymerase also revealed the same result. Jalview 2.8 [25] was used to visualize the result. This tool also provides a bar diagram of conservation summery. Yellow colour bar with star sign indicates full conservation. Consensus sequences are shown in black bar and yellow colour indicates good quality. Depends on conservation and alignment quality colours are changed.

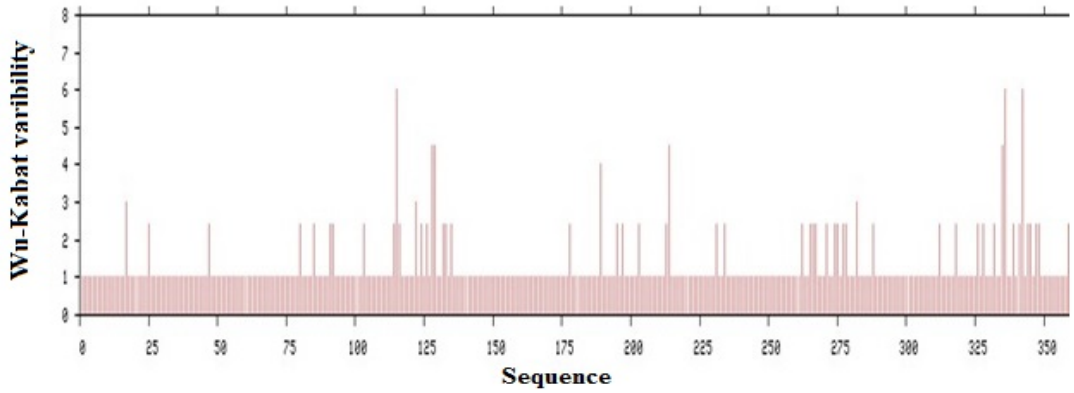

Fig. (2). Protein variability index of the conserved peptide. The conservancy threshold was 1.0 in this analysis. The X-axis indicates the amino acid positions in the sequences and the Y-axis indicates Wu-Kabat variability score.

Table 1. Antigenicity determination of conserved peptide by VaxiJen server.

Peptide number	Peptide	Region	Length	VaxiJen Score (Threshold: 0.5)
1	RLAWTIGNRAPYIGSRTEDKIGYP-PLRVNCPSAALKEAIEMVSRLLWVTQGTADREKL-LIPLLNSRVNLDYQTVLNFLPTHYSGNIVHRYNDQYGQHSF-MANRMSNTSTRAIISTNTLGKYAGG GQAA	1207-1334	127	**0.6752**
2	FAEKLEYLAPSYRNFSFSLKEKELNIGRT-FGKLPYRVRNVQTLAEALLADGLAKAFPSNMMVVTEREQ-KEALLHQASWHHNSASIGENAIVR-GASFVTDLEKYNLAFRYEFTRHFI	537-652	115	**0.6346**

Table 2. T-cell epitope prediction by NetCTL server.

Number	Sequence	Combined Score (nM)
1	YSGNIVHRY	2.9568
2	WTIGNRAPY	2.4674
3	YRVRNVQTL	2.4425
4	APSYRNFSF	1.8800
5	FRYEFTRHF	1.8309

Table 3. Prediction of the T-cell epitope by CTL Pred server.

Peptide Rank	Start Position	Sequence	Score(ANN/SVM)
1	135	YLAPSYRNF	0.87/0.98139506
2	220	RGASFVTDL	0.96/0.74768804
3	8	NRAPYIGSR	0.94/0.69872428
4	125	GQAAFAEKL	0.96/0.65920328
5	1	RLAWTIGNR	0.96/0.6437403

Table 4. MHC-Iand MHC-II interaction of the proposed sequence by IEDB-AR.

	Epitope	MHC Interaction
	MHC-I Interaction Analysis	
1	YSGNIVHRY	HLA-C*03:03, HLA-C*14:02, HLA-C*12:03, HLA-C*07:02, HLA-A*31:01, HLA-A*29:02, HLA-C*07:01, HLA-A*68:01, HLA-A*01:01, HLA-C*05:01, HLA-C*15:02
2	WTIGNRAPY	HLA-C*03:03, HLA-C*12:03, HLA-A*31:01, HLA-A*29:02, HLA-A*26:01, HLA-A*30:02, HLA-C*07:01, HLA-C*07:02, HLA-C*14:02, HLA-B*35:01, HLA-B*15:01, HLA-B*15:02, HLA-C*15:02, HLA-C*07:01
3	YRVRNVQTL	HLA-C*03:03, HLA-C*07:02, HLA-B*15:02, HLA-C*14:02, HLA-C*12:03, HLA-B*39:01, HLA-C*07:01, HLA-C*06:02, HLA-B*27:05, HLA-A*30:01, HLA-C*15:02
4	APSYRNFSF	HLA-C*12:03, HLA-C*03:03, HLA-B*15:02, HLA-C*07:02, HLA-C*14:02, HLA-B*15:01, HLA-B*35:01, HLA-C*07:01, HLA-B*27:05, HLA-A*30:01, HLA-C*15:02
5	FRYEFTRHF	HLA-C*06:02, HLA-C*12:03, HLA-A*68:01, HLA-C*07:02, HLA-C*07:01, HLA-A*31:01, HLA-B*15:02, HLA-C*03:03, HLA-C*14:02, HLA-B*27:05, HLA-B*40:02, HLA-B*18:01, HLA-B*40:01, HLA-A*23:01, HLA-B*44:03, HLA-A*14:02

Table 4. cond…

	Epitope	MHC Interaction
	MHC-II interaction analysis	
1	LAWTIGNRAPYIGSR	HLA-DRB1*13:02, HLA-DRB1*03:01, HLA-DRB1*09:01, HLA-DRB1*08:02, HLA-DRB1*11:01, HLA-DRB3*02:02, HLA-DRB1*15:01, HLA-DRB1*12:01, HLA-DRB1*07:01, HLA-DQA1*01:01/DQB1*05:01, HLA-DRB1*04:01, HLA-DRB3*01:01, HLA-DRB5*01:01, HLA-DQA1*05:01/DQB1*03:01, HLA-DRB1*04:05, HLA-DQA1*03:01 /DQB1*03:02, HLA-DRB1*01:01, HLA-DPA1*03:01/DPB1*04:02, HLA-DPA1*02:01/DPB1*14:01, HLA-DQA1*05:01/DQB1*02:01, HLA-DQA1*01:02/DQB1*06:02, HLA-DPB1*01:01, HLA-DQA1*04:01/DQB1*04:02, HLA-DPA1*01:03/DPB1*02:01, HLA-DQA1*01:02/DQB1*06:02, HLA-DPB*04:01,HLA-DPB*05:01, HLA-DRB4*01:01.

Table 5. Bepipred and IEDB analysis tools provided B-cell epitope and their lengths.

No.	Bepipred Analysis	Length
1.	RAPYIGSRTEDKIGY	15
2.	GTADR	5
3.	NDQYG	5
4.	LGKYAGGGQAA	11
5.	IGSYDFP	7
6.	LRVN	4
7.	SNT	3
8.	REQ	3
9.	ASIG	4

Table 6. Kolasker and Tongaonkar antigenicity prediction.

No.	Peptide	Start	End	Length
1	GYPPLRVNCPSAALKE	22	37	16
2	IEMVSRLLWVTQG	39	51	13
3	EKLLIPLLNSR	56	66	11
4	NLDYQTVLNFLPTH	68	81	14
5	GNIVHRY	84	90	7
6	NFSFSL	142	147	6
7	RAPYIGS	9	15	7
8	VRGASFVTDL	219	228	10
9	EALLHQAS	198	205	8
10	KYNLAFRY	230	237	8
11	AAFAEKLEYLAPSY	127	140	14
12	KLPYRVRNVQTLAEALLADGLAKAF	160	184	25

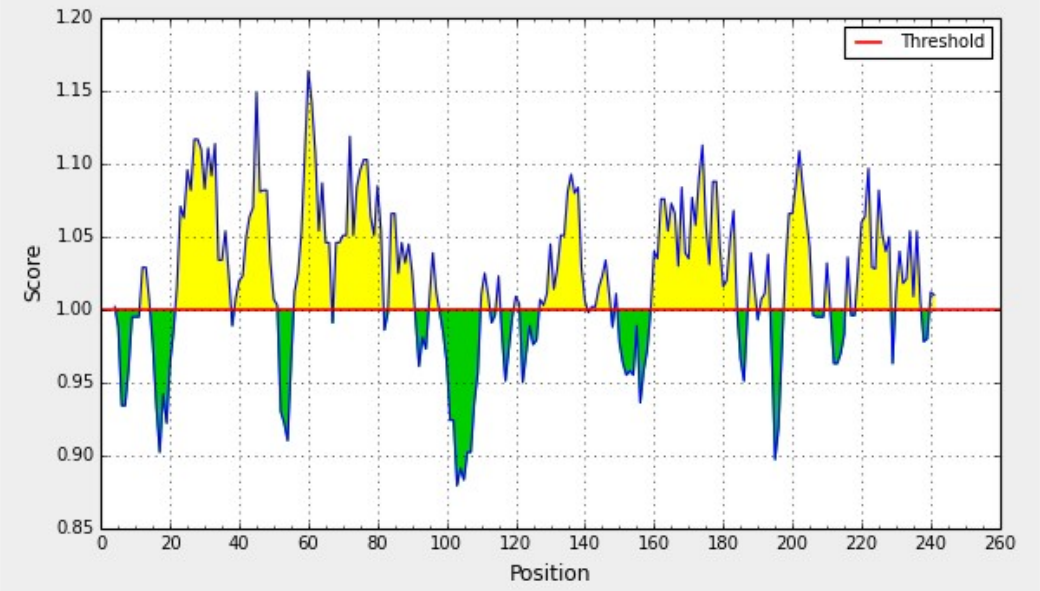

Fig. (3). Antigenicity of the conserved peptide. The X- and Y-axes represent the sequence position and antigenic propensity score, respectively. The threshold value is 1.0. The regions above the threshold are antigenic, shown in yellow.

Table 7. Emini surface accessibility prediction.

No.	Peptide	Start	End	Length
1.	GNRAPY	7	12	6
2.	SRTEDKIG	15	22	8
3.	GTADRE	51	56	6
4.	RYNDQYG	89	95	7
5.	RMSNTS	103	108	6
6.	LPYRVR	161	166	6
7.	TEREQK	192	197	6
8.	DLNKYMD	252	258	7

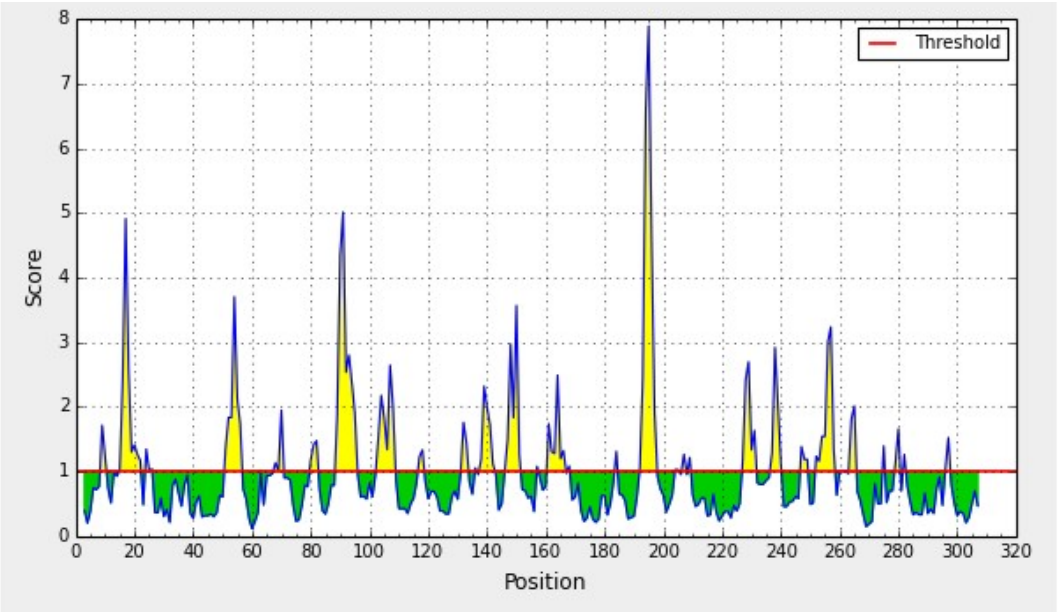

Fig. (4). The surface accessible residues of the conserved peptide which are above the cut off are located in the yellow region. The red horizontal line indicates surface accessibility cut off (1.000).

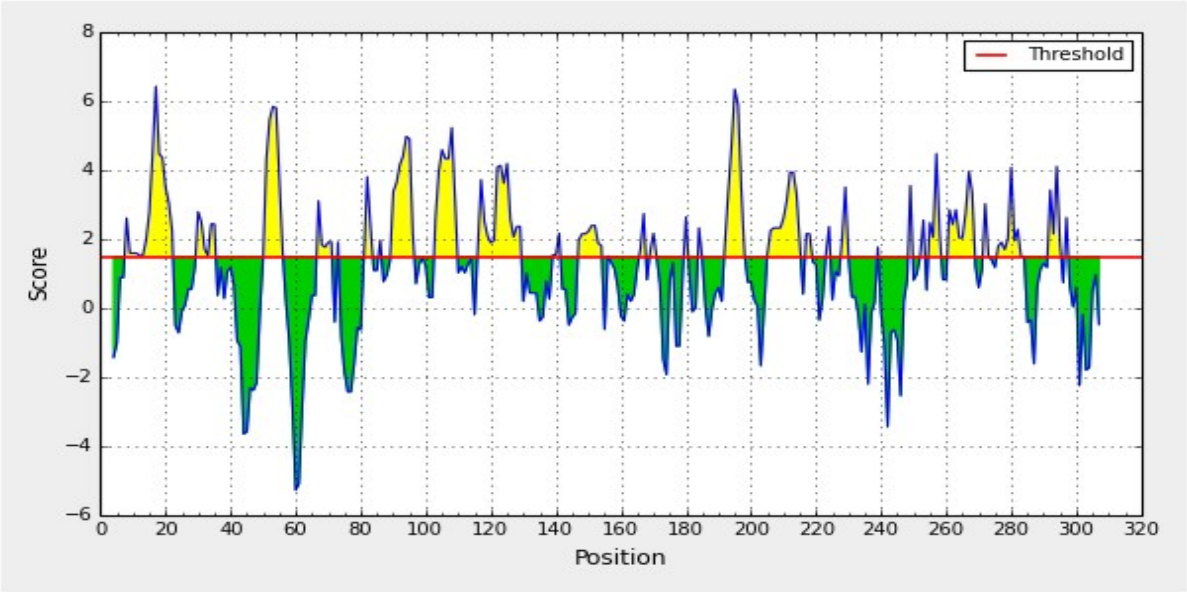

Fig. (5). IEDB analysis provided antigenicity score of the conserved peptide. Most of the residues were found above the threshold 1.00 values. Residues in the yellow region are antigenic and in the green regions are below the threshold (red line).

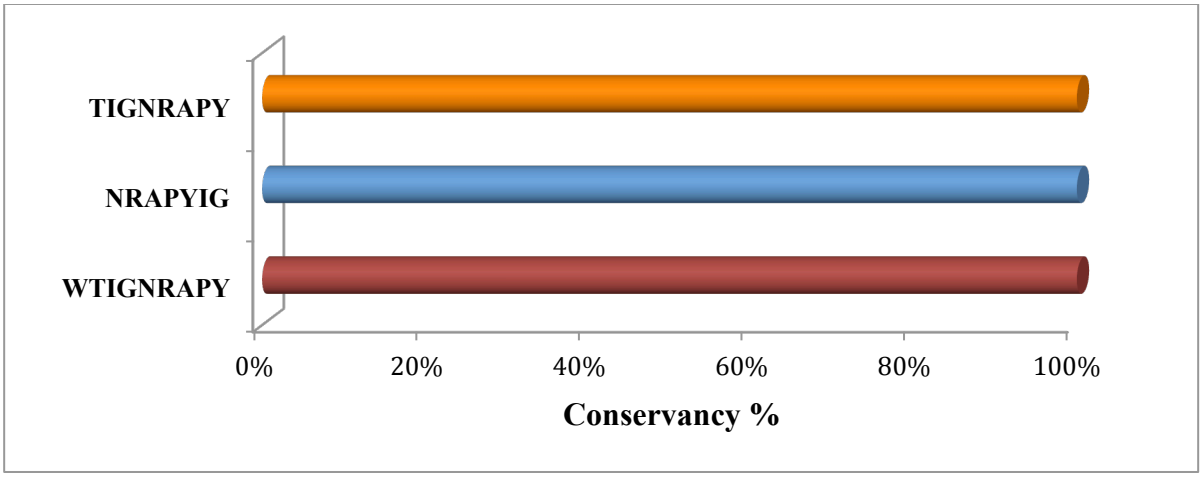

Fig. (6). Conservancy percentages of the predicted consensus epitopes. All of the three epitopes were found 100% conserved. X-axis indicates the conservancy percentage and Y-axis indicates the epitopes.

Out of these eight epitopes, GNRAPY epitope overlaps with the Bepipred [38] and Kolaskar and Tongaonkar antigenicity [40] predicted common epitope NRAPYIG. Accessibility of the epitope can be determined by the property of hydrophilicity of the epitope. Hydrophilicity was determined by the IEDB Parker hydrophilicity [42] prediction tool at threshold level 1.5. The maximum level was found as 6.443 and minimum as -5.286 in the epitope (Fig. **5**).

3.5. TIGNRAPYI is Thoroughly Conserved Among all Human Marburg Virus Strains and Non-Allergenic

Epitope conservancies were identified by IEDB conservancy [44] analysis tool. NetCTL and CTLpred predicted common epitope WTIGNRAPY was found to be 100% conserved. Bepipred and IEDB analyzed common epitope NRAPYIG was found to be 100% conserved. This WTIGNRAPY and NRAPYIG epitopes overlap each other and 100% conserved in all human Marburg viruses. So the potential general vaccine candidate epitope was chosen as TIGNRAPYI, which was also 100% conserved (Fig. **6**).

Allergenicity of the proposed epitope was determined by AlgPred [43] based on amino acid composition. The proposed epitope was found to be non-allergenic. The prediction score was 0.65674 at threshold level-0.4.

3.6. Conserved sequence 3D structure identification and analysis

Because of the inaccessibility of the 3D structure of RNA directed RNA polymerase, we modelled the 3D structure of the conserved region by I-TASSER [45] server. I-TASSER analysis depicts 5 different models for this sequence by multiple threading alignments. Stereochemical quality of all the predicted models was tested by PROCHECK [46] server through the Ramachandran Plot. The desired best model was

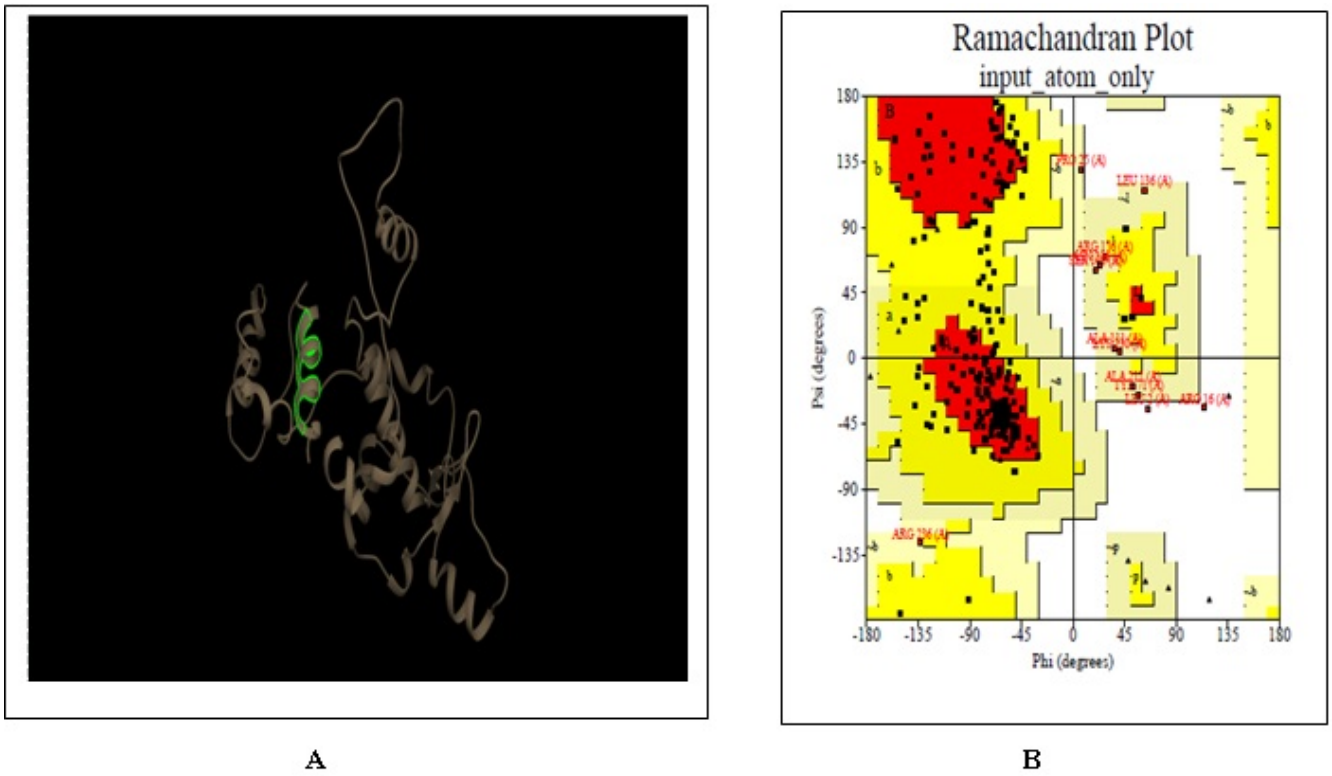

Fig. (7). 3D structure prediction and validation. (**A**) 3D structure of RNA directed RNA polymerase protein. The conserve TIGNRAPYI epitope is highlighted using UCSF Chimera visualization tool on the structure. The green colour indicates the position of conserve TIGNRAPYI epitope. (**B**) The validity of the model was proved by the Ramachandran plot of the predicted model, which showed that most of the residues are in the favoured region of the plot.

selected from PROCHECK analysis results, based on the factors that have maximum numbers of amino acids residues in the favorable region and highest G factor. The potential best model is given in Fig. **7A**. In the best model, the G factor was -0.50 and 91.4% residues are found in the favored region in the Ramachandran plot (Fig. **7B**). TIGNRAPYI epitope was also found in surface and accessible in the model.

4. DISCUSSION

Marburg virus is one of the most harmful infectious viruses and day by day its high fatality rate that makes this virus as a deadly virus. All the seven strains which are identified around the world are so deadly and some of them caused a serious outbreak in several areas in recent times [12]. According to the WHO, the virus is rated as Risk Group 4 Pathogen [48]. Both Marburg and Ebola viruses comprise the same taxonomic family and both viruses have similar physiological structure although they produce different antibodies. So vaccine design against this virus is very important at this time. Epitope-based vaccine development is widely accepted because of its specificity and it shows no side effects. It also stimulates an effective specific immune response within a minimal structure [49]. A conserved protein sequence like glycoprotein, RNA-directed RNA polymerases, nucleoprotein or any other membranes which are unchanged during the lifetime is needed to design a peptide vaccine. To design an appropriate peptide vaccine against Marburg virus, we tried to find a conserved sequence. We did multiple sequence alignment of the glycoproteins, RNA-directed RNA polymerases, Nucleoproteins, Vp24 proteins, Vp30, Vp35, and Vp40 protein sequences of all strain of Marburg virus in search of conserved sequence. Except for RNA-directed RNA polymerases protein, we did not find any conserved sequence from other protein. But we found a conserved region among all the available strains of Marburg virus in RNA-directed RNA polymerases protein.

Then the RNA directed RNA polymerase protein was selected to observe its activity against T-cell and B-cell, whether it activates T-cell and B-cell immunity with a single epitope or not. Antigenicity and surface accessibility were calculated using different bioinformatics analyses. Antigenicity score was intended to determine T-cell epitopes. The epitope TIGNRAPYI showed the best result among all epitopes in activating T-cell immunity with potential antigenicity score in both NetCTL and CTLPred output analyses. In MHC-I and MHC-II binding analysis, the epitope TIGNRAPYI showed the highest number of binding interaction with the most number of HLA loci. As a result, MHC-I molecules would easily present the epitope on the T-cell surface. On the other hand in case of MHC-II interaction analysis, TIGNRAPYI is included in LAW-TIGNRAPYIGSR. The epitope TIGNRAPYI also has the capability to induce B-cell immune responses and we confirm this by amino acid-based B-cell epitope prediction analysis which comprises Kolaskar and Tongaonkar antigenicity scale, Emini surface accessibility prediction, and Bepipred linear epitope prediction analysis. The accessibility and hydrophilicity property of this epitope makes it a perfect candidate to be used as a vaccine. The epitope TIGNRAPYI was found to be 100% conserved among all its strains. This

epitope is also found as nonallergic. As the three-dimensional structure of the conserved peptide was not available on the database, we modelled a three-dimensional structure of the conserved peptide of the protein (Figure 7A). Ramachandran Plot referred that 91.4% of residues are within the favourable region, which indicates the validation of the proposed model. An effective vaccine shows high immunogenicity and stability inside the body. As the epitope-based vaccine is designed by *in silico* analysis process, it needs *in vitro* and *in vivo* experimental analysis, because the real immunogenicity, stability, efficacy of the epitope cannot be determined by this *in silico* analysis. *In vitro*, this epitope-based vaccine has been verified to work [50]. So based on the above analyses, we can expect that our suggested vaccine will elicit *in vitro* immune responses.

CONCLUSION

Our study shows that conserved sequences among same virus genus but different strains can be used to design an epitope-based universal vaccine against some deadly viruses such as Marburg virus. So, our designed epitope would be a good therapeutic agent for all human Marburg virus strains. The wet-laboratory experiment is needed to resolve the definite effect of the peptide for inducing an immune response. Thus computational studies help the researchers to get their desired results with greater feasibility.

ABBREVIATIONS

ANN/SVM = ...artificial neural networks/support vector machines

CTLP = ...Cytotoxic T Lymphocyte Prediction (CTLPred)

MSA = ... multiple-sequence alignment

IEDB-AR = ...Immune Epitope Database and Analysis Resource

MHC = ...major histocompatibility complex.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable.

HUMAN AND ANIMAL RIGHTS

No Animals/Humans were used for studies that are the basis of this research.

CONSENT FOR PUBLICATION

Not applicable.

AVAILABILITY OF DATA AND MATERIALS

Not applicable.

FUNDING

None.

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

Declared none.

REFERENCES

[1] Kuhn, J.H.; Becker, S.; Ebihara, H.; Geisbert, T.W.; Johnson, K.M.; Kawaoka, Y.; Lipkin, W.I.; Negredo, A.I.; Netesov, S.V.; Nichol, S.T.; Palacios, G.; Peters, C.J.; Tenorio, A.; Volchkov, V.E.; Jahrling, P.B. Proposal for a revised taxonomy of the family Filoviridae: classification, names of taxa and viruses, and virus abbreviations. *Arch. Virol.,* **2010**, *155*(12), 2083-2103. [http://dx.doi.org/10.1007/s00705-010-0814-x] [PMID: 21046175]

[2] Kiley, M.P.; Bowen, E.T.; Eddy, G.A.; Isaäcson, M.; Johnson, K.M.; McCormick, J.B.; Murphy, F.A.; Pattyn, S.R.; Peters, D.; Prozesky, O.W.; Regnery, R.L.; Simpson, D.I.; Slenczka, W.; Sureau, P.; van der Groen, G.; Webb, P.A.; Wulff, H. Filoviridae: a taxonomic home for Marburg and Ebola viruses? *Intervirology,* **1982**, *18*(1-2), 24-32. [http://dx.doi.org/10.1159/000149300] [PMID: 7118520]

[3] Geisbert, T.W.; Jahrling, P.B. Differentiation of filoviruses by electron microscopy. *Virus Res.,* **1995**, *39*(2-3), 129-150. [http://dx.doi.org/10.1016/0168-1702(95)00080-1] [PMID: 8837880]

[4] Mühlberger, E.; Lötfering, B.; Klenk, H.D.; Becker, S. Three of the four nucleocapsid proteins of Marburg virus, NP, VP35, and L, are sufficient to mediate replication and transcription of Marburg virus-specific monocistronic minigenomes. *J. Virol.,* **1998**, *72*(11), 8756-8764. [PMID: 9765419]

[5] Becker, S.; Huppertz, S.; Klenk, H.D.; Feldmann, H. The nucleoprotein of Marburg virus is phosphorylated. *J. Gen. Virol.,* **1994**, *75*(Pt 4), 809-818. [http://dx.doi.org/10.1099/0022-1317-75-4-809] [PMID: 8151297]

[6] Modrof, J.; Möritz, C.; Kolesnikova, L.; Konakova, T.; Hartlieb, B.; Randolf, A.; Mühlberger, E.; Becker, S. Phosphorylation of Marburg virus VP30 at serines 40 and 42 is critical for its interaction with NP inclusions. *Virology,* **2001**, *287*(1), 171-182. [http://dx.doi.org/10.1006/viro.2001.1027] [PMID: 11504552]

[7] Becker, S.; Rinne, C.; Hofsäss, U.; Klenk, H.D.; Mühlberger, E. Interactions of Marburg virus nucleocapsid proteins. *Virology,* **1998**, *249*(2), 406-417. [http://dx.doi.org/10.1006/viro.1998.9328] [PMID: 9791031]

[8] Siegert, R.; Shu, H.L.; Slenczka, W.; Peters, D.; Müller, G. Zur Atiologie einer umbekannten, von Affen ausgegangenen menschlichen Infektionskrankheit. *Dtsch. Med. Wochenschr.,* **1967**, *92*(51), 2341-2343. [http://dx.doi.org/10.1055/s-0028-1106144] [PMID: 4294540]

[9] US Animal and Plant Health Inspection Service (APHIS) and US Centers for Disease Control and Prevention (CDC), National Select Agent Registry (NSAR). **2011**, 10-16.

[10] Martini, GA *Siegert Reds.,* **1971**.

[11] Funke, C.; Becker, S.; Dartsch, H.; Klenk, H.D.; Mühlberger, E. Acylation of the Marburg virus glycoprotein. *Virology,* **1995**, *208*(1), 289-297. [http://dx.doi.org/10.1006/viro.1995.1151] [PMID: 11831710]

[12] Gear, J.S.; Cassel, G.A.; Gear, A.J.; Trappler, B.; Clausen, L.; Meyers, A.M.; Kew, M.C.; Bothwell, T.H.; Sher, R.; Miller, G.B.; Schneider, J.; Koornhof, H.J.; Gomperts, E.D.; Isaäcson, M.; Gear, J.H. Outbreake of Marburg virus disease in Johannesburg. *BMJ,* **1975**, *4*(5995), 489-493. [http://dx.doi.org/10.1136/bmj.4.5995.489] [PMID: 811315]

[13] Gear, J.H. Haemorrhagic fevers of Africa: an account of two recent outbreaks. *J. S. Afr. Vet. Assoc.,* **1977**, *48*(1), 5-8. [PMID: 406394]

[14] Bausch, D.G.; Borchert, M.; Grein, T.; Roth, C.; Swanepoel, R.; Libande, M.L.; Talarmin, A.; Bertherat, E.; Muyembe-Tamfum, J.J.; Tugume, B.; Colebunders, R.; Kondé, K.M.; Pirad, P.; Olinda, L.L.; Rodier, G.R.; Campbell, P.; Tomori, O.; Ksiazek, T.G.; Rollin, P.E. Risk factors for Marburg hemorrhagic fever, Democratic Republic of the Congo. *Emerg. Infect. Dis.,* **2003**, *9*(12), 1531-1537. [http://dx.doi.org/10.3201/eid0912.030355] [PMID: 14720391]

[15] Beer, B.; Kurth, R.; Bukreyev, A. Characteristics of Filoviridae: Marburg and Ebola viruses. *Naturwissenschaften,* **1999**, *86*(1), 8-17.

[http://dx.doi.org/10.1007/s001140050562] [PMID: 10024977]

[16] Hovette, P. *Epidemic of Marburg hemorrhagic fever in Angola.,* **2005**.

[17] Timen, A.; Koopmans, M.P.; Vossen, A.C.; van Doornum, G.J.; Günther, S.; van den Berkmortel, F.; Verduin, K.M.; Dittrich, S.; Emmerich, P.; Osterhaus, A.D.M.E.; van Dissel, J.T.; Coutinho, R.A. Response to imported case of Marburg hemorrhagic fever, the Netherland. *Emerg. Infect. Dis.,* **2009**, *15*(8), 1171-1175. [http://dx.doi.org/10.3201/eid1508.090015] [PMID: 19751577]

[18] Wu, X.; Smith, T.G.; Rupprecht, C.E. From brain passage to cell adaptation: the road of human rabies vaccine development. *Expert Rev. Vaccines,* **2011**, *10*(11), 1597-1608. [http://dx.doi.org/10.1586/erv.11.140] [PMID: 22043958]

[19] Plotkin, S.A. Vaccines: past, present and future. *Nat. Med.,* **2005**, *11*(4)(Suppl.), S5-S11. [http://dx.doi.org/10.1038/nm1209] [PMID: 15812490]

[20] Islam, R.; Sadman, S.; Aubhishek, Z. A computational assay to design an epitope-based peptide vaccine against Chikungunya virus. *Future Virol.,* **2012**, *7*, 1029-1042. [http://dx.doi.org/10.2217/fvl.12.95]

[21] Sette, A.; Fikes, J. Epitope-based vaccines: an update on epitope identification, vaccine design and delivery. *Curr. Opin. Immunol.,* **2003**, *15*(4), 461-470. [http://dx.doi.org/10.1016/S0952-7915(03)00083-9] [PMID: 12900280]

[22] Poland, G.A.; Ovsyannikova, I.G.; Jacobson, R.M. Application of pharmacogenomics to vaccines. *Pharmacogenomics,* **2009**, *10*(5), 837-852. [http://dx.doi.org/10.2217/pgs.09.25] [PMID: 19450131]

[23] Petrovsky, N.; Brusic, V. Computational immunology: The coming of age. *Immunol. Cell Biol.,* **2002**, *80*(3), 248-254. [http://dx.doi.org/10.1046/j.1440-1711.2002.01093.x] [PMID: 12067412]

[24] Holland, J.; Domingo, E. Origin and evolution of viruses. *Virus Genes,* **1998**, *16*(1), 13-21. [http://dx.doi.org/10.1023/A:1007989407305] [PMID: 9562888]

[25] Sette, A.; Newman, M.; Livingston, B.; McKinney, D.; Sidney, J.; Ishioka, G.; Tangri, S.; Alexander, J.; Fikes, J.; Chesnut, R. Optimizing vaccine design for cellular processing, MHC binding and TCR recognition. *Tissue Antigens,* **2002**, *59*(6), 443-451. [http://dx.doi.org/10.1034/j.1399-0039.2002.590601.x] [PMID: 12445314]

[26] Benson, D.A.; Cavanaugh, M.; Clark, K.; Karsch-Mizrachi, I.; Lipman, D.J.; Ostell, J.; Sayers, E.W. GenBank. *Nucleic Acids Res.,* **2013**, *41*(Database issue), D36-D42. [PMID: 23193287]

[27] Thompson, J.D.; Higgins, D.G.; Gibson, T.J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.,* **1994**, *22*(22), 4673-4680. [http://dx.doi.org/10.1093/nar/22.22.4673] [PMID: 7984417]

[28] Waterhouse, A.M.; Procter, J.B.; Martin, D.M.; Clamp, M.; Barton, G.J. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics,* **2009**, *25*(9), 1189-1191. [http://dx.doi.org/10.1093/bioinformatics/btp033] [PMID: 19151095]

[29] Garcia-Boronat, M; Diez-Rivero, CM; Reinherz, EL; Pedro, AR *PVS: a web server for protein sequence variability analysis tuned to facilitate conserved epitope discovery.,* **2008**. [http://dx.doi.org/10.1093/nar/gkn211]

[30] Doytchinova, I.A.; Flower, D.R. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinformatics,* **2007**, *8*, 4-8. [http://dx.doi.org/10.1186/1471-2105-8-4] [PMID: 17207271]

[31] Larsen, M.V.; Lundegaard, C.; Lamberth, K.; Buus, S.; Lund, O.; Nielsen, M. Large-scale validation of methods for cytotoxic T-lymphocyte epitope prediction. *BMC Bioinformatics,* **2007**, *8*, 424-436. [http://dx.doi.org/10.1186/1471-2105-8-424] [PMID: 17973982]

[32] Bhasin, M.; Raghava, G.P. Prediction of CTL epitopes using QM, SVM and ANN techniques. *Vaccine,* **2004**, *22*(23-24), 3195-3204. [http://dx.doi.org/10.1016/j.vaccine.2004.02.005] [PMID: 15297074]

[33] Buus, S.; Lauemøller, S.L.; Worning, P.; Kesmir, C.; Frimurer, T.; Corbet, S.; Fomsgaard, A.; Hilden, J.; Holm, A.; Brunak, S. Sensitive quantitative predictions of peptide-MHC binding by a 'Query by Committee' artificial neural network approach. *Tissue Antigens,* **2003**, *62*(5), 378-384. [http://dx.doi.org/10.1034/j.1399-0039.2003.00112.x] [PMID: 14617044]

[34] Wang, P.; Sidney, J.; Kim, Y.; Sette, A.; Lund, O.; Nielsen, M.; Peters, B. Peptide binding predictions for HLA DR, DP and DQ molecules. *BMC Bioinformatics,* **2010**, *11*, 568-580. [http://dx.doi.org/10.1186/1471-2105-11-568] [PMID: 21092157]

[35] Wang, P.; Sidney, J.; Dow, C.; Mothé, B.; Sette, A.; Peters, B. A systematic assessment of MHC class II peptide binding predictions and evaluation of a consensus approach. *PLOS Comput. Biol.,* **2008**, *4*(4)e1000048 [http://dx.doi.org/10.1371/journal.pcbi.1000048] [PMID: 18389056]

[36] Peters, B.; Sette, A. Generating quantitative models describing the sequence specificity of biological processes with the stabilized matrix method. *BMC Bioinformatics,* **2005**, *6*, 132. [http://dx.doi.org/10.1186/1471-2105-6-132] [PMID: 15927070]

[37] Nair, D.T.; Singh, K.; Siddiqui, Z.; Nayak, B.P.; Rao, K.V.; Salunke, D.M. Epitope recognition by diverse antibodies suggests conformational convergence in an antibody response. *J. Immunol.,* **2002**, *168*(5), 2371-2382. [http://dx.doi.org/10.4049/jimmunol.168.5.2371] [PMID: 11859128]

[38] Larsen, J.E.; Lund, O.; Nielsen, M.; Morten, N. Improved method for predicting linear B-cell epitopes. *Immunome Res.,* **2006**, *2*, 2-9. [http://dx.doi.org/10.1186/1745-7580-2-2] [PMID: 16635264]

[39] Vita, R.; Zarebski, L.; Greenbaum, J.A.; Emami, H.; Hoof, I.; Salimi, N.; Damle, R.; Sette, A.; Peters, B. The immune epitope database 2.0. *Nucleic Acids Res.,* **2010**, *38*(Database issue), D854-D862. [http://dx.doi.org/10.1093/nar/gkp1004] [PMID: 19906713]

[40] Kolaskar, A.S.; Tongaonkar, P.C. A semi-empirical method for prediction of antigenic determinants on protein antigens. *FEBS Lett.,* **1990**, *276*(1-2), 172-174. [http://dx.doi.org/10.1016/0014-5793(90)80535-Q] [PMID: 1702393]

[41] Emini, E.A.; Hughes, J.V.; Perlow, D.S.; Boger, J. Induction of hepatitis A virus-neutralizing antibody by a virus-specific synthetic peptide. *J. Virol.,* **1985**, *55*(3), 836-839. [PMID: 2991600]

[42] Parker, J.M.; Guo, D.; Hodges, R.S. New hydrophilicity scale derived from high-performance liquid chromatography peptide retention data: correlation of predicted surface residues with antigenicity and X-ray-derived accessible sites. *Biochemistry,* **1986**, *25*(19), 5425-5432. [http://dx.doi.org/10.1021/bi00367a013] [PMID: 2430611]

[43] Saha, S.; Raghava, G.P. AlgPred: prediction of allergenic proteins and mapping of IgE epitopes. *Nucleic Acids Res.,* **2006**, *34*(Web Server issue)(Suppl. 2)W202-9 [http://dx.doi.org/10.1093/nar/gkl343] [PMID: 16844994]

[44] Bui, H.H.; Sidney, J.; Li, W.; Fusseder, N.; Sette, A. Development of an epitope conservancy analysis tool to facilitate the design of epitope-based diagnostics and vaccines. *BMC Bioinformatics,* **2007**, *8*, 361. [http://dx.doi.org/10.1186/1471-2105-8-361] [PMID: 17897458]

[45] Zhang, Y. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics,* **2008**, *9*, 40. [http://dx.doi.org/10.1186/1471-2105-9-40] [PMID: 18215316]

[46] Laskowski, R.A.; MacArthur, M.W.; Moss, D.S.; Thornton, J.M. PROCHECK: a program to check the stereochemical qualit of protein structures. *J. Appl. Cryst.,* **2001**, *26*, 283-291. [http://dx.doi.org/10.1107/S0021889892009944]

[47] Pettersen, E.F.; Goddard, T.D.; Huang, C.C.; Couch, G.S.; Greenblatt, D.M.; Meng, E.C.; Ferrin, T.E. UCSF Chimera--a visualization system for exploratory research and analysis. *J. Comput. Chem.,* **2004**, *25*(13), 1605-1612. [http://dx.doi.org/10.1002/jcc.20084] [PMID: 15264254]

[48] US Department of Health and Human Services. Biosafety in Microbiological and Biomedical Laboratories (BMBL) 5th Edition **2011**, 10-16.

[49] Arnon, R. A novel approach to vaccine design–epitope-based vaccines. *FEBS J.,* **2006**, *273*, 33-34.

[50] Khan, M.K.; Zaman, S.; Chakraborty, S.; Chakravorty, R.; Alam, M.M.; Bhuiyan, T.R.; Rahman, M.J.; Fernández, C.; Qadri, F.; Se-

raj, Z.I. *In silico* predicted mycobacterial epitope elicits *in vitro* T-cell responses. *Mol. Immunol.,* **2014**, *61*(1), 16-22. [http://dx.doi.org/10.1016/j.molimm.2014.04.009] [PMID: 24853589]

[51] Hossain, M.U.; Keya, C.A.; Das, K.C.; Hashem, A.; Omar, T.M.; Khan, M.A.; Rakib-Uz-Zaman, S.M.; Salimullah, M.; Salimullah, M. An immunopharmacoinformatics approach in development of vaccine and drug candidates for West Nile virus. *Front Chem.,* **2018**, *6*, 246. [http://dx.doi.org/10.3389/fchem.2018.00246] [PMID: 30035107]

[52] Hasan, M.A.; Khan, M.A.; Datta, A.; Mazumder, M.H.H.; Hossain, M.U. A comprehensive immunoinformatics and target site study revealed the corner-stone toward Chikungunya virus treatment. *Mol. Immunol.,* **2015**, *65*(1), 189-204. [http://dx.doi.org/10.1016/j.molimm.2014.12.013] [PMID: 25682054]

[53] Khan, M.A.; Hossain, M.U.; Rakib-Uz-Zaman, S.M.; Morshed, M.N. Epitope-based peptide vaccine design and target site depiction against Ebola viruses: an immunoinformatics study. *Scand. J. Immunol.,* **2015**, *82*(1), 25-34. [http://dx.doi.org/10.1111/sji.12302] [PMID: 25857850]

[54] Sharmin, R.; Islam, A.B. A highly conserved WDYPKCDRA epitope in the RNA directed RNA polymerase of human coronaviruses can be used as epitope-based universal vaccine design. *BMC Bioinformatics,* **2014**, *15*, 161. [http://dx.doi.org/10.1186/1471-2105-15-161] [PMID: 24884408]