We are IntechOpen, the world’s leading publisher of Open Access books
Built by scientists, for scientists

4,000
Open access books available

116,000
International authors and editors

120M
Downloads

154
Countries delivered to

TOP 1%
Our authors are among the most cited scientists

12.2%
Contributors from top 500 universities

WEB OF SCIENCE™
Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com
Predicting Candidate Epitopes on Ebola Virus for Possible Vaccine Development

Maryam Hemmati, Ehsan Raoufi and Hossein Fallahi

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.72413

Abstract

Zaire ebolavirus, a member of family Filoviridae is the cause of hemorrhagic fever. Due to lack of appropriate antiviral or vaccine, this disease is very lethal. In this study, we tried to find epitopes for superficial glycoprotein and nucleoprotein of Zaire ebolavirus (that have high antigenicity for MHC I, II and B cells) by using in silico methods and immunoinformatics approach. By using CTLPred, SYFPEITHI and ProPred web applications for MHC class I and SYFPEITHI and ProPred1 web applications for MHC class II, we had been able to find epitopes (peptides) that have the highest score. Also ElliPro, IgPred and DiscoTope web tools had been performed to predict B cells conformational epitopes. Linear epitope prediction for B cell was performed with six methods from IEDB. All of the results that including candidate epitopes for T cells and B cells were reported. It was expected that these peptides could be stimulated immune response and used for designing the multipeptide vaccine against ZEV but these results should be reliable with experimental analysis.

Keywords: epitopes, glycoprotein, immune response, immunoinformatics approach, multipeptide vaccine, nucleoprotein, Zaire ebolavirus

1. Introduction

Zaire ebolavirus, a member of genus Ebola virus, family Filoviridae, and order Mononegavirales, is enveloped, RNA negative strand genome and filamentous virus. This virus is the cause of serious hemorrhagic fever (HF) in human and the mortality rate is 50–90% [1, 2]. There have been several outbreaks of Ebola virus up to now and the last one was on 2014. This outbreak started in Guinea and spread to other countries in West Africa. The 2014 outbreak was the largest one which had the highest mortality rate and risk of spreading to different parts of the world [3, 4].
Experimental work with EBOV is very difficult and currently, there is no effective and licensed vaccine available for Ebola. Vaccination is a good approach for the prevention and treatment against many types of diseases like the viral infection. There are several types of vaccines, and the goal of all of them is presenting antigenic fragment to the immune system to induce the adaptive immune response. Several types of antiviral vaccines consist of inactivated, live attenuated, virus-like particles and DNA vaccines. Another type of vaccine is peptide-based vaccines. These vaccines are based on designing different epitopes for B cells and T cells. These vaccines, in comparison with other vaccines, have fewer side effects and are safe and easy to prepare [4–6]. Obviously, T cells have an important role in stimulating immune responses, but for having a better and accurate response, first of all, antigenic fragments should be attached to major histocompatibility complex (MHC) molecules. MHC molecules process and present antigenic peptides to T cells. These peptide epitopes must be linear for attaching to MHC molecules [7]. The most of T cells are belonging to two groups namely CD8+ and CD4+. The difference between two groups refers to different glycoproteins in the surface of T cells. CD8+ T cells are cytotoxic T cells (CTL) that bind to MHC class I molecules and CD4+ T cells are T helper cells that attach to MHC class II [8, 9]. B cells are another part of immune systems, which have receptors and secrete antibodies. Epitopes can be discontinuous or continuous for B cells. These epitopes can also bind to lipids, carbohydrates and peptides but 3D structure of antigens has an important role in stimulating humoral immune response by B cells [10]. With the advances in in silico method and bioinformatics, immunoinformatics, a branch of science was progressed. Immunoinformatics is an interdisciplinary science that emanates from immunology and bioinformatics and generates meaningful immunological data. With the help of immunoinformatics approaches, we can find epitopes for B cells and T cells, and also we can design vaccines based on peptide or multipeptide vaccine [4, 11].

In this chapter, we tried to use immunoinformatics tools and in silico method to predict MHCs linear epitopes and B cells discontinuous and linear epitopes (peptides) for glycoprotein and nucleoprotein of Zaire ebolavirus (as antigen). These results can be useful for finding epitopes that can be the candidate for designing vaccine and therapeutic strategies for fighting with HF (hemorrhagic fever) of Ebola virus.

2. The best candidate for producing vaccines against Ebola virus

The ZEOBV genome has seven ORF, including NP-VP35-VP40-GP-VP30-VP24-L. Nucleo-protein (NP) causes encapsulation of the genome of Ebola. It has been aggregated with VP30 (transcription factor) and VP35 (polymerase cofactor). L is RNA-dependent RNA polymerase. VP24 is a minor matrix protein that associates with the membrane. VP40 is a major matrix protein that can mediate virus particles creation [1, 4, 12]. Glycoprotein (GP) is present on the surface of Ebola. Secreted nonstructural GP, structural GP 1 and GP 2 are the results of mRNA editing during transcription of GP of ZEBOV. The N-terminal sequence between the GP and sGP is the same but the C-terminal of them is different [4, 13]. The ratio between sGP and GP during infection is 80–20%. Superficial GP is cleaved by furin in Golgi to GP1 and GP2 that form homotrimetric proteins on the surface of EBOV.
The GP1 is a soluble protein on the surface of virus, which has a mucin-like domain and is highly glycosylated. GP2 is a membrane-spanning subunit and smaller than GP1, which is connected to GP1 by disulfide bonds [14]. GP has a major role in attaching to host cells and has cytotoxicity effects. This polycistronic GP is the main difference between ZEBOV and other mononegavirus. Because of the GP on the surface, this protein has the most antigenicity for the immune system and so it can be the best candidate for producing vaccines against EBOV [2, 4]. Immunoinformatic studies on Ebola virus proteins also have shown that there is a high rate of immune response for epitopes of Ebola nucleoprotein (NP), in which, it can be even more than the response for epitopes of Ebola glycoprotein [15–18]. Therefore, it can be another target for peptide-based vaccine designing against Ebola virus.

Bioinformatics evaluation on the superficial glycoprotein showed less than 30% identity with family and less than 50% with species of identity in the this protein, this glycoprotein is also glycosylated. Studies have been shown that the structure and sequencing of GP are constantly changing and these changes are the major causes of the weakness of the immune system against the virus and, consequently, its pathogenicity. For these reasons, designing and development of the vaccine against the Ebola virus is difficult and complex.

3. Epitope mapping for the purpose of peptide-based vaccine design

The most commonly used assimilation and the concept of the term epitope mapping is the mapping of the antigenic regions detected by antibodies. This term is ambiguous because the purpose of using this term in research is not clear, and this mapping can be done with a variety of goals. Therefore, researchers have tried to use alternative terms for epitope mapping in articles in order to clarify the purpose of this bioinformatics process such as immunological analysis, microstructure analysis and epitope determination. Despite all these efforts and including a wide range of goals, scientists are still advised to use this term in articles and resources.

A number of important applications of epitope mapping including:

- Determination of biological process mechanism,
- Recognition of an epitope of practical value,
- Connection of any type of polymorphism such as SNP to protein structure or Ab binding,
- Characterization of Ab binding in patients,
- Evaluation of vaccine design,
- Identification of autoimmune diseases,
- Qualification of an Ab for diagnostic use such as Western blot analysis, trans-species assays, finding isoforms of Ag and allergen characterization,
- Distinguishing of antigen peptide mimic,
Establishing of antigen structure and

Finding an antigen which it can make different between immunization antibodies and infection antibodies.

Immunogenicity studies have shown epitope mapping (the branch of bioinformatics) is based on mathematical modeling and evolutionary algorithms. Epitope mapping methods can be used to molecular modeling behaviors in nature, so the data obtained from these methods can predict molecular interactions, such as binding of antigen to antibody and peptide to MHC, with high probability [19, 20].

Designing of vaccines based on peptides or epitopes is the goal that we have been considering in this study. They are within the epitope mapping compass and to achieve these goals, we identify and predict the epitopic areas by using approved and high-performance software.

Mathematically modeled methods for epitope mapping are artificial neural networks (ANN), quantization matrices, decision trees, HMM (Markov secret models), SMM is a stabilized matrix method (SMM), among which ANN, SVM and HMM are capable of analyzing linear and nonlinear data. In many applications that analyze epitope prediction, these three methods are used to identify the sequential linear epitopes of the T lymphocytes and the nonlinear (spatial) epitopes of the B lymphocytes [4, 21, 22].

There are other predictive methods for detecting spatial epitopes of the B cell lymphocytes, which include the following: homology modeling, docking, 3D- and threading techniques. These methods with the ANN, SMM and HMM methods are most used in the study of the spatial epitope and are among the software main methods used in this research [4, 23]. In sum, all methods for identifying epitopes that have the best antigenic properties but the epitopes should have important characteristics, including (1) structural flexibility, (2) in the surface of the protein, (3) exposure to solvent, (4) containing charged amino acids and (5) contains hydrophilic amino acids.

4. Candidate for designing peptide-based vaccines against Ebolavirus

Conventional vaccines are prepared from an attenuated or inactive version of the pathogen. However, often the antigen to which the immune system responds is of small number of amino acids or peptide of antigen. The alternative approach to stimulate the immune response, such as humoral and cell-mediated immune response, is the identification of peptide sequences or epitopes that have a protective immune response. Therefore, these epitopes can have no risk of mutation and they would be more stable, and also, there can be fewer side effects [24, 25]. This approach is called peptide-based vaccine designing.

4.1. Determine the T cell antigenic fragment

Epitope analysis was performed for MHC classes I and II.
4.1.1. Epitopes analysis for MHC class I

By using CTLPred (http://www.imtech.res.in/raghava/ctlpred/index.html), SYFPEITHI (http://www.syfpeithi.de/bin/MHCServer.dll/EpitopePrediction.htm) and ProPred1 (http://www.imtech.res.in/raghava/propred/) web tools, we had been able to find some peptidic epitopes for MHC class I.

A. Finding epitopes for MHC class I by using CTLPred web tool and sequence of GP and NP of ZEBOV. In Tables 1 and 2, three peptides that have higher score are shown for each protein. Positions 19, 245 and 130 for GP and positions 292, 263 [4] and 397 for NP have higher scores, so these peptides have been suggested by the application to be epitopes.

B. Using SYFPEITHI web tool for predicting epitopes. We investigated some alleles of the MHC-I for the protein of interest. According to these results, as is shown in Table 3, for GP 564, 246 and 205 positions have a higher score in this prediction web tools [4]. But when we investigated other alleles, we had observed that the position 246 was the most repetitive of all and had a high score. For NP, results are shown in Table 4, position 266 has the highest score.

C. ProPred1 was used to predict epitopes of GP and NP, which bind with the highest score to MHC class I alleles. The peptide with highest score was selected and is shown in

<table>
<thead>
<tr>
<th>Strat position</th>
<th>Sequence</th>
<th>Score(ANN\SVM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>130</td>
<td>RGFPRCRYV</td>
<td>0.97/0.99110928</td>
</tr>
<tr>
<td>245</td>
<td>ESRFTPQFL</td>
<td>0.98/0.867525</td>
</tr>
<tr>
<td>19</td>
<td>FFLWVIIIF</td>
<td>0.97/078863896</td>
</tr>
</tbody>
</table>

Table 1. Epitope prediction of GP based on CTLPred web server.

<table>
<thead>
<tr>
<th>Strat position</th>
<th>Sequence</th>
<th>Score(ANN\SVM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>292</td>
<td>EYAPFARLL</td>
<td>0.96/1.5915619</td>
</tr>
<tr>
<td>263</td>
<td>RLHPLARTA</td>
<td>0.89/1.2678105</td>
</tr>
<tr>
<td>397</td>
<td>LRRERLAKL</td>
<td>0.84/1.0652382</td>
</tr>
</tbody>
</table>

Table 2. Epitope prediction of NP based on CTLPred web server.
Tables 5 and 6. According to this result, the highest score belonged to position 264 for the HLA-B*2705 allele for GP, and as shown in Table 6 for NP, positions 273 and 109 for HLA-A20 allele have higher scores.

Finally with revision of all data that are achieved for MHC class I, we are able to conclude that sequences: “TRKIRSEEL” (with position 298 for T residue), “SRFTPQFLL” (with position 246 for S residue) and “ETTQLQLF” (with position 561 for E residue) for GP [4], and sequences: “RLHPLARTA” (with position 263 for first residue), “SRELDHLGL” (with position 360 for...
first residue) and “VKNEVNSFK” (with position 273 for first residue), have the highest score and the most frequent within these analysis. Therefore, it can be a better candidate peptide epitopes than other sequences.

4.1.2. Epitopes analysis for MHC class II

By using SYFPEITHI and ProPred web tools, some peptidic epitopes for MHC class II were predicted.
A. Using SYFPEITHI web tools, epitope mapping of the amino acid sequence of the GP and NP of ZEBOV was performed. Accordingly, we investigated some alleles of MHC class II for the protein of interest. Peptides that have higher score were selected and are shown in Table 7 for GP and in Table 8 for NP. According to this result for GP, the highest score belongs to position 135 with the score of 35 and after that, the position of 159 has a higher score.

---

**Table 5.** Epitope prediction from GP for MHC class I by ProPred1 web tools (only top score peptides are shown).

<table>
<thead>
<tr>
<th>Allele</th>
<th>Sequence</th>
<th>Position</th>
<th>Real score</th>
</tr>
</thead>
<tbody>
<tr>
<td>HLA-B*2705</td>
<td>TRKIRSEEL</td>
<td>298</td>
<td>10000</td>
</tr>
<tr>
<td>HLA-B*2705</td>
<td>GREAAVSHL</td>
<td>357</td>
<td>9000</td>
</tr>
<tr>
<td>HLA-B*2705</td>
<td>FQRTFSIPL</td>
<td>27</td>
<td>6000</td>
</tr>
</tbody>
</table>

**Table 6.** NP epitope prediction of MHC class I according to ProPred1 web server.

<table>
<thead>
<tr>
<th>Allele</th>
<th>Sequence</th>
<th>Position</th>
<th>score</th>
</tr>
</thead>
<tbody>
<tr>
<td>HLA-A20 cattle</td>
<td>VKNEVNSFK</td>
<td>273</td>
<td>4000</td>
</tr>
<tr>
<td>HLA-A20 cattle</td>
<td>KKRDGVKRL</td>
<td>109</td>
<td>4000</td>
</tr>
<tr>
<td>HLA-B2705</td>
<td>SRELDHLGL</td>
<td>360</td>
<td>3000</td>
</tr>
</tbody>
</table>

**Table 7.** GP epitope prediction for MHC class II based on SYFPEITHI web server.

<table>
<thead>
<tr>
<th>Allele</th>
<th>Sequence</th>
<th>Position</th>
<th>score</th>
</tr>
</thead>
<tbody>
<tr>
<td>HLA_DRB1*0101</td>
<td>CRYHKVSSTGPGCAG</td>
<td>135</td>
<td>35</td>
</tr>
<tr>
<td>HLA_DRB1*0101</td>
<td>FFLYDRLASTVIYRG</td>
<td>159</td>
<td>33</td>
</tr>
<tr>
<td>HLA_DRB1*0701</td>
<td>ILFQRTFSIPLGVI</td>
<td>24</td>
<td>34</td>
</tr>
<tr>
<td>HLA_DRB1*0101</td>
<td>LRTFSILNRKADFL</td>
<td>579</td>
<td>30</td>
</tr>
</tbody>
</table>
score for the HLA-DRB1*0101 allele [4]. For NP, position 710 with score 38 have the highest score.

B. ProPred web tool was used to predict epitopes. Peptide with the highest score was selected and is shown in Table 9 for GP and in Table 10 for NP. This result for GP has

<table>
<thead>
<tr>
<th>Allele</th>
<th>Sequence</th>
<th>Position</th>
<th>score</th>
</tr>
</thead>
<tbody>
<tr>
<td>DRB1*0703</td>
<td>FQRTFSIPL</td>
<td>26</td>
<td>9.7000</td>
</tr>
<tr>
<td>DRB1*0817</td>
<td>FFLYDRLAS</td>
<td>158</td>
<td>7.2000</td>
</tr>
<tr>
<td>DRB1*0701</td>
<td>FQRTFSIPL</td>
<td>26</td>
<td>9.7000</td>
</tr>
</tbody>
</table>

Table 9. Epitope prediction of GP for MHC class II based on ProPred web server.
been shown that the positions 158 and 26 have the highest score. These positions are very frequent in reviewing other results. For NP positions, 149 and 200 have higher score according to ProPred web server.

As a result for MHC class II for GP, sequences: “IILFQRTFSIPLGVI” (with position 24 for first residue), “CRYVHKVSGTGPCAG” (with position 135 for the first residue) and “FFLYDTLAS” (with position 158 for first the residue) [4] and for NP sequences: “FLSFASLFL” (with position 149 for the first residue), “NRFVTLDGQQFYWPV” (with position 710 for first residue) and “FRLMRTNFL” (with position 200 for first residue) have the highest scores and the most frequent within these analysis.

By considering these results for MHC classes I and II, we think these epitopes can activate the cell-mediated immune response; therefore they can be used for producing peptide-based vaccines.

4.2. Determine the B cell antigenic fragment

4.2.1. Prediction of linear (sequential) epitopes

In this section, six methods from IEDB were used for the prediction of linear epitopes. A collection of methods to predict linear B cell epitopes based on sequence characteristics of the antigen using amino acid scales and HMMs. These methods include the following:
1. BepiPred Linear Epitope Prediction: BepiPred predicts the location of linear B cell epitopes using a combination of a HMM and a propensity scale method [26].

2. Chou & Fasman Beta-Turn Prediction: This method is commonly used to predict beta turns to the prediction of antibody epitopes.

3. Emini Surface Accessibility Prediction: The computation was based on surface accessibility scale on antibodies. The accessibility profile was achieved using the formulae $S_n = (n + 4 + i) (0.37)^{-6}$, where $S_n$ is the surface probability, $d_n$ is the fractional surface probability value, and $i$ vary from 1 to 6. A hexapeptide sequence with $S_n$ greater than 1.0 indicates an increased probability of existing on the surface [27].

4. Karplus & Schulz Flexibility Prediction: In this method, flexibility scale based on ability to move protein segments. The calculation based on a flexibility scale is similar to classical calculation, except that the center is the first amino acid of the six amino acids window length, and there are three scales for describing flexibility instead of a single one [28].

5. Kolaskar & Tongaonkar Antigenicity: A semiempirical method that makes use of physicochemical properties of amino acid residues and their frequencies of occurrence in experimentally known segmental epitopes was developed to predict antigenic peptide established on proteins. This method can predict antigenic peptide with about 75% accuracy [29].

6. Parker Hydrophilicity Prediction: In this method, hydrophilic scale based on peptide retention times during high-performance liquid chromatography (HPLC) on a reversed-phase column was formulated [30].

4.2.1.1. Linear epitopes prediction for GP

We performed all of six methods for the GP protein sequence and summarized its results in Table 11.

<table>
<thead>
<tr>
<th>Method number</th>
<th>Threshold</th>
<th>Max score</th>
<th>The best candidate peptide sequences</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.271</td>
<td>2.51 for 384 position</td>
<td>370 to 394</td>
</tr>
<tr>
<td>2</td>
<td>1.007</td>
<td>1.437 for 384 position</td>
<td>381 to 387</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>7.532 for 499 position</td>
<td>Scattered areas</td>
</tr>
<tr>
<td>4</td>
<td>1.008</td>
<td>1.119 for 321 position</td>
<td>318 to 324</td>
</tr>
<tr>
<td>5</td>
<td>1.015</td>
<td>1.225 for 183 position</td>
<td>180 to 186</td>
</tr>
<tr>
<td>6</td>
<td>1.793</td>
<td>7.957 for 640 position</td>
<td>637 to 643</td>
</tr>
</tbody>
</table>

Table 11. Conclusion results for B cell linear epitope prediction from GP protein sequence.
4.2.1.2. Linear epitopes prediction for NP

We performed all of six methods for the NP protein sequence and summarized its results in Table 12.

4.2.2. Prediction of discontinuous (conformational) epitopes

4.2.2.1. Discontinuous epitopes prediction for GP

For B cell epitope prediction, 3D structure of antigen is more important than linear sequence, therefore, in this study, we used PDB ID of GP of ZEBOV.

By considering the structure of GP in PDB, we comprehended that “I” chain of protein has a maximum length than other chains. Also, with investigating two epitopes that predicted for MHCs molecules, we understood both of them are on the “I” chain. Therefore, this chain was selected in this analysis.

A. Using ElliPro web tool, epitope mapping of the amino acid sequence of the GP of ZEBOV was performed. PDB Id (3csy) and “I” chain was used for ElliPro tool. Results from this prediction for linear epitopes have been shown that the sequence of the peptides from 31 to 64 and 255 to 310 had a higher score than the other part of the protein and are illustrated in Figures 1–3 and Table 13 [4].

B. By using DiscoTope and 3D structure of GP of ZEV, we could find discontinuous epitopes. These sequences may be near to each other in 3D conformation but far from each other in the amino acid sequence or the first structure. In this study, we analyze the “I” chain from GP1 and we set the threshold on –7.7, upstream regions of the threshold have the positive prediction. Two regions that are shown in Figure 4 have positive predictions but 261–310 regions have a more score and these scores for each amino acid are illustrated in Table 14 [4].

In Table 14, contact number indicates the number of amino acids that are next to each amino acid. As much as the contact number is lower, it shows that our amino acids are

<table>
<thead>
<tr>
<th>Method number</th>
<th>Threshold</th>
<th>Max score</th>
<th>The best candidate peptide sequences</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.359</td>
<td>2.777 for 477 position</td>
<td>Scattered areas</td>
</tr>
<tr>
<td>2</td>
<td>0.991</td>
<td>1.511 for 437 position</td>
<td>434 to 440</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>5.323 for 371 position</td>
<td>369 to 374</td>
</tr>
<tr>
<td>4</td>
<td>1.007</td>
<td>1.141 for 507 and 508 positions</td>
<td>504 to 511</td>
</tr>
<tr>
<td>5</td>
<td>1.016</td>
<td>1.22 for 43 position</td>
<td>40 to 46</td>
</tr>
<tr>
<td>6</td>
<td>2.101</td>
<td>8.686 for 492 position</td>
<td>489 to 495</td>
</tr>
</tbody>
</table>

Table 12. Conclusion results for B cell linear epitope prediction from NP protein sequence.
more external. For example, in this study, VAL residue in the 310th position had the lowest contact number and was more external.

C. Using IgPred web tool, epitope mapping of the amino acid sequence of the GP of ZEBOV was performed. The sequence of the chain “I” of glycoprotein was selected to study the interaction with the antibodies IgG, IgE and IgA.

As illustrated in Table 15, different regions in this chain had a different score for IgG but the end of the sequence had the highest score and was approximately 255–310 regions.
Therefore, according to prediction with these web tools, 255–310 regions are the proper candidate for being epitope [4].

4.2.2.2. Discontinuous epitopes prediction for NP

No specific structure was found in the PDB for NP. Therefore, homology modeling of these types of proteins is needed to determine their structures. This goal has been achieved with the help of homology modeling, and its stages have been described below:

A. **Template Selection:** The template to be used in homology modeling should be based on target-template alignment, and the template that most closely resembles with our protein is selected as the template.

B. **Model Building:** The models are constructed using target-template alignment by ProMod3, and the areas that are shared and conserved between the template and the target are copied from the template on the model. The areas that are added or removed are rebuilt using fragment library. Side chains are then remodeled. Finally, its overall geometry is determined, but the loop areas are building with PROMOD-II.

---

Table 13. Three top scores predicted discontinuous epitope for B cell according to ElliPro web server.

<table>
<thead>
<tr>
<th>No.</th>
<th>Residues</th>
<th>Number of residues</th>
<th>Score</th>
</tr>
</thead>
</table>

---

Figure 3. Discontinuous epitope(s) number three 3D structure for GP.
Table 14. The predicted epitope sequence from 261 to 310 discontinuous amino acids for B cell according to DiscoTope web server.

<table>
<thead>
<tr>
<th>Chain ID</th>
<th>Residue ID</th>
<th>Residue Name</th>
<th>Contact number</th>
<th>DiscoTope Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>261</td>
<td>TYR</td>
<td>13</td>
<td>-7.447</td>
</tr>
<tr>
<td>I</td>
<td>262</td>
<td>THR</td>
<td>11</td>
<td>-5.510</td>
</tr>
<tr>
<td>I</td>
<td>263</td>
<td>SER</td>
<td>11</td>
<td>-5.510</td>
</tr>
<tr>
<td>I</td>
<td>264</td>
<td>GLY</td>
<td>9</td>
<td>-2.952</td>
</tr>
<tr>
<td>I</td>
<td>265</td>
<td>LYS</td>
<td>15</td>
<td>-6.372</td>
</tr>
<tr>
<td>I</td>
<td>267</td>
<td>SER</td>
<td>15</td>
<td>-7.348</td>
</tr>
<tr>
<td>I</td>
<td>268</td>
<td>ASN</td>
<td>10</td>
<td>-3.780</td>
</tr>
<tr>
<td>I</td>
<td>269</td>
<td>THR</td>
<td>10</td>
<td>-4.544</td>
</tr>
<tr>
<td>I</td>
<td>270</td>
<td>THR</td>
<td>8</td>
<td>-3.423</td>
</tr>
<tr>
<td>I</td>
<td>271</td>
<td>GLY</td>
<td>13</td>
<td>-6.966</td>
</tr>
<tr>
<td>I</td>
<td>277</td>
<td>VAL</td>
<td>10</td>
<td>-5.284</td>
</tr>
<tr>
<td>I</td>
<td>278</td>
<td>ASN</td>
<td>8</td>
<td>-3.870</td>
</tr>
<tr>
<td>I</td>
<td>299</td>
<td>ARG</td>
<td>7</td>
<td>-2.755</td>
</tr>
<tr>
<td>I</td>
<td>300</td>
<td>LYS</td>
<td>5</td>
<td>-0.883</td>
</tr>
<tr>
<td>I</td>
<td>301</td>
<td>ILE</td>
<td>11</td>
<td>-5.702</td>
</tr>
<tr>
<td>I</td>
<td>302</td>
<td>ARG</td>
<td>8</td>
<td>-2.804</td>
</tr>
<tr>
<td>I</td>
<td>303</td>
<td>SER</td>
<td>12</td>
<td>-4.847</td>
</tr>
<tr>
<td>I</td>
<td>304</td>
<td>GLU</td>
<td>10</td>
<td>-4.323</td>
</tr>
<tr>
<td>I</td>
<td>306</td>
<td>LEU</td>
<td>11</td>
<td>-7.687</td>
</tr>
<tr>
<td>I</td>
<td>309</td>
<td>THR</td>
<td>7</td>
<td>-6.575</td>
</tr>
<tr>
<td>I</td>
<td>310</td>
<td>VAL</td>
<td>3</td>
<td>-3.807</td>
</tr>
</tbody>
</table>

Figure 4. The model 3D structure.
C. **Results:** The SWISS-MODEL template library (SMTL version 2017-09-21, PDB release 2017-09-15) was searched with BLAST and HHBlits for evolutionarily related structures matching the target sequence is presented in Table 16 and the model 3D structure is also shown in Figure 4.

D. After building homology modeling, ElliPro web tool was performed to predict conformational epitopes of NP. These results are shown in Table 17 and (Figures 5 and 6).

### Table 16. Details on the template search.

<table>
<thead>
<tr>
<th>Template</th>
<th>Seq Identity</th>
<th>Oligo-state</th>
<th>Found by</th>
<th>Method Resolution</th>
<th>Seq Similarity</th>
<th>Range</th>
<th>Coverage</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>4yp1_2_A</td>
<td>100.00</td>
<td>hetero-oligomer</td>
<td>HHBlits</td>
<td>X-ray 3.71Å</td>
<td>0.60</td>
<td>39-384</td>
<td>0.47</td>
<td>Nucleoprotein</td>
</tr>
</tbody>
</table>

### Table 17. Top scores predicted discontinuous epitope for B cell according to ElliPro web server.

<table>
<thead>
<tr>
<th>No.</th>
<th>Scores</th>
<th>Number of residues</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4</td>
<td>71</td>
<td>4.179</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>4</td>
<td>4.179</td>
</tr>
</tbody>
</table>

---

Table 15. Antibodies scores for sequence of “I” chain of GP according to IgPred web server.
5. Conclusion

The hemorrhagic fever (HF) is a lethal disease from ZEBOV that caused the death of many people in Africa. The glycoprotein of ZEBOV is the only protein on the surface of the virus, and has severe cytotoxicity effects, also nucleoprotein can stimulate the immune response, stronger than GP, therefore, it was offered to design the vaccine against both of them [4, 13,
Vaccination is a good idea to prevent infection and limit the spreading of Ebola virus. So far, several vaccines against Ebola glycoprotein have been tested on animals, but no licensed vaccine has been reported in humans [31, 32]. There are several types of vaccine, in which one of them is a peptide vaccine, and it is the goal of this analysis. These vaccines are designed based on the epitopes fragment of antigens and they are safer and easy to prepare. For T cells, these epitopes are linear and they must bind to MHC I and MHC II at first, but for B cells, 3D conformation of antigens is very important so it can be discontinuous epitopes [6]. These discontinuous epitopes may be near to each other in the 3D conformation but far from each other in the amino acid sequence or first structure [4, 10]. The experimental work on this virus is very difficult and biosafety level is 4. With the help of immunoinformatics tools for vaccine design, the use of laboratory work is reduced, and hence we save more expense and time. Immunoinformatics approach can help us to predict T cell and B cell epitopes, and also have application in in silico vaccination. By predicting T cell and B cell epitopes, we can find the peptide that is useful for in silico vaccination or for designing multipeptide vaccine [11]. By using ProPred, CTLPred, SYFPEITHI and ProPred1 web tools for predicting the epitopes for T cells, we had been able to introduce peptides that can be the candidate for designing multipeptide vaccines. With the use of ElliPro, DiscoTope, IgPred web servers and linear epitope prediction methods from IEDB predicting epitopes for B cells, and these peptides can induce the immune response and design for the peptide-based vaccine. With the help of immunoinformatics tools for predicting epitopes for T cells and B cells, we can design the multipeptide vaccine, and this vaccine can include both epitopes from GP and NP, which is useful for increasing immune response against Ebola virus. As conclusion for linear epitopes that bind to MHC I “TRKIRSEEL,” “SRFTPQFLL” and “ETTQLALQLF” peptides for GP [4], and “RLHPLARTA,” “SRELDHGLL” and “VKNEVNSFK” peptides for NP are candidate epitopes. Also for MHC class II: “ILLFQRTFSPLGVI,” “CRYVHKVSGTGPCAG” and “FFLYDTLAS” peptides for GP and “FLSFASLFL,” “NRFVTLDGQQFYWPV” and “FRLMRTNFL” peptides for NP have the highest scores and the most frequent within this analysis. For B cell linear epitope prediction, results are shown in Tables 11 and 12. As a final result for conformational epitopes, we can only say, peptide sequence from 255 to 310 amino acids for GP has a higher score. These peptides are able to be the candidate for the vaccine against Ebola virus. It should be noted that these results just in in silico are valid and need laboratory (in vivo and in vitro) confirmation.

Author details

Maryam Hemmati†, Ehsan Raoufi‡*† and Hossein Fallahi‡

*Address all correspondence to: ehsan.raoufi@gmail.com

1 Department of Medical Biotechnology, School of Allied medicine, Iran University of Medical Sciences, Tehran, Iran

2 Department of Biology, Faculty of Science, Razi University, Kermanshah, Iran

† These authors contributed equally.
References


