Epitope Prediction and Structural Analysis of Sterol 24-c-Methyltransferase Antigen of Leishmania donovani Using In Silico Approach

Manju Kashyap1, Azhar Khan2 and Umar Farooq1,3*

1Immunoparasitology Laboratory, Faculty of Applied Sciences and Biotechnology, Shoolini University, Solan, India
2Faculty of Dentistry, Taif University, Taif, Saudi Arabia
3School of Computer Sciences, Shoolini University, Solan, India

*Corresponding Author: Umar Farooq, Professor, Microbiology, Faculty of Applied Sciences and Biotechnology, Shoolini University, Solan, HP, India.

Abstract

Visceral leishmaniasis is an important vector born infectious disease. It is caused by the protozoan parasites of genus Leishmania including Leishmania donovani and Leishmania infantum. Although, there are chemotherapeutic treatments against leishmaniasis, but none of them are completely effective. In silico methods have generated the new dimension in peptide based vaccine development. The present study was done to identify promiscuous T-cell epitopes for MHC class-II alleles and B-cell epitopes targeting sterol 24-c-methyltransferase protein using in silico approaches. The identified T and B cell epitopes were confirmed by visualizing their locations on the respective 3D modelled protein. The binding pattern of the identified peptides with predominant HLA class II alleles (HLA*DRB1 0101, DRB1*1501) was also provided. The identified T cell and B-cell epitopes may help to know insights in better understanding the humoral and adaptive immune responses generated in response to vaccine candidate (SMT) protein of the parasite. It can help in the development of subunit vaccine.

Keywords: Leishmania; T-cell; B-cell; Epitope; HLA; Alleles

Introduction

Leishmaniasis is a vector born disease and prevalent in 98 countries worldwide [1]. Leishmania parasite causes several human infections ranging from Moderate form of disease, Cutaneous and Mucocutaneous leishmaniasis to severe form of Visceral leishmaniasis (VL). Other animals including dogs, rodents and marsupials are also affected from the Leishmania parasite [2]. It is estimated that about 2 million new cases of leishmaniasis and 59,000 deaths occur annually [3]. Organisms mainly causing VL are Leishmania donovani and Leishmania infantum. There are several chemotherapeutic agents used to treat VL including pentavalent antimonials, amphotericin B and its liposomal formulations and miltefosine to control the disease in human beings [4]. Although, chemotherapeutic agents have some drawbacks including toxicity, painful administration, higher cost and besides all these, emerging and spread of drug resistance is an alarming feature [5]. There is a need to develop an efficient, safe and cost effective adjuvant to overcome problems linked with the treatment of visceral leishmaniasis. Vaccine is an important alternative to control the disease. Vaccine if develop will be a safe and cost intensive adjuvant. The identification and characterization of leishmanial antigens, which is able to elicit protective immune response and important to develop a subunit vaccine. The characterization of a given antigen in a laboratory is a tedious, time consuming and cost intensive process. Recent development in bioinformatics (Immuninformatics) is an alternative method, which is a rapid and cost effective method to characterize the antigens to develop a sub-unit vaccine [6]. By using immuninformatics tool, we can screen the published gene sequences to identify the B and T cell epitopes, which may recognise polymorphic HLA alleles.

A major challenge in epitope based vaccine development is to identify the immunogenic sites of antigenic proteins, which can generate greatest immune response [7]. But, using immuninformatics methods, sequence regions in antigenic proteins with potential binding sites for both T-cell and B-cell epitopes can be identified and peptides of varying lengths can be synthesized [8]. T-cell epitopes are linear peptides and able to bind to both HLA class-I and class-II alleles. The identification and analysis of HLA binding within the antigenic proteins is important for vaccine development [9]. It is cleared that protective immune response in the visceral leishmaniasis is based on the generation of Th1 mediated immune response [10]. T-helper cells generate the adaptive immunity after parasite infection due to the recognition of epitopes of antigens presented by HLA class-II alleles [11-13]. Therefore, play an important role in elimination of Leishma-
Epitope Prediction and Structural Analysis of Sterol 24-c-Methyltransferase Antigen of Leishmania donovani Using In Silico Approach

The present study sterol 24-c-methyltransferase (SMT) protein of *L. donovani* was taken for analysis. It is an important vaccine candidate antigen. This enzyme plays an important role in the biosynthesis of ergosterol, which is act as a target for leishmanicidal and fungicidal amphotericin B [16]. The earlier studies have reported that the SMT protein based vaccine protect challenge [3]. The present study was aimed to identify B cell and T cell epitopes by screening of SMT protein by using in-silico methods.

**Material and Methods**

**Retrieval of sequence of protein**

The amino acid sequence of SMT protein of *Leishmania donovani* was retrieved from the National Centre for Biotechnology (NCBI). The protein sequence was saved in FASTA format.

**Secondary structure prediction**

The prediction of secondary structure of SMT was done by using improved self optimized prediction method (SOPMA) software ([http://npsa-pbil.ibcp.fr/cgibin/npsa_automat.pl?page=/NPSA/npsa_sopma.html](http://npsa-pbil.ibcp.fr/cgibin/npsa_automat.pl?page=/NPSA/npsa_sopma.html)) [17]. The parameters related to similarity threshold and window width were set to 8 and 17, remaining parameters were not adjusted.

**Prediction of linear B-cell epitopes**

Identification of T-cell epitopes for HLA class-II was done by selecting the class-II alleles: HLA-DRB1*0101, HLA-DRB1*0701, HLA-DRB1*1101, HLADRBI*1301, HLA-DRB1*1401, HLA-DRB1*1501. The servers utilized for prediction of these alleles were: NetMHCIpan2.1 [18], NetMHCIIpan3.0 [19], ED8-ANN, SMM (Stabilized Matrix alignment Method), ANN (Artificial Neural Network) algorithms, ARB (Average Relative Binding) [20]. T-cell epitopes were predicted based on their binding affinity for HLA alleles using half maximal inhibitory concentration of a biological substance (IC50). Strong binders showed IC50 value < 500 nm and non-binder peptides showed IC50 value > 5000 nm [21]. Strong binders were taken for analysis and non-binders were not considered for further analysis.

**Results**

**Protein sequence retrieval**

The complete amino acid sequence of SMT antigen was obtained from NCBI database having accession number: AAR92098.1. This protein was screened for the prediction of epitopes. This protein is also found to be conserved in various species of *Leishmania*. The protein sequence was saved in FASTA format.

**Secondary structure prediction**

The antigenic behaviour of the SMT was accessed by predicting the secondary structure with the help of SOPMA server. It is studied that proteins representing greater proportion of extended strands and random coils are able to form antigenic epitopes (Li, et al. 2013). The results of predicted structure are demonstrated in figure 1.

Selected promiscuous T-cell epitopes

For the selection of promising T-cell epitopes, it should have least IC$_{50}$ value i.e. < 500 nm towards the all HLA-DRB1 alleles (DRB1*0101, DRB1*0701, DRB1*1101, DRB1*1301, DRB1*1401, DRB1*1501). Only peptides, which were having 100% binding affinity with all alleles, were selected as promiscuous in nature. Only two promiscous (CRVLEFVRLAPKGTY and RVLEFVRLAPKGTYK) peptides were identified as strong binders from SMT antigen.

Homology modeling of HLA-DRB1 alleles and T-cell epitopes

The 3D structures of alleles and peptides were prepared by homology modeling. First of all templates were identified. As alpha chain is similar in all HLA DRB1 alleles, PDB blast of beta chain of alleles was run (DRB1*0701, HLA DRB1*1301, HLA DRB1*1401, HLA DRB1*1101). For DRB1*0701, 1AQD was used, for DRB1*1101, 2SEB was used, for DRB1*1301 and DRB1*1401, 1A6A was used as template. The other alleles DRB1*0101 and DRB1*1501 were have solved crystal structures in RCSB database. Modeling of alleles was performed by using Modeller9v7 program, all models were optimized and minimized to provide stability. Similarly, UCSF Chimera 1.8.1 was used to prepare 3D models of the promiscuous peptides [25] and energy minimization was done for providing stability.

Verification of structures

The modelled alleles were further verified by Ramachandran plot and ERRAT. The residues of HLA-DRB1*0701, DRB1*1101, DRB1*1301 and DRB1*1401 allele showed 93.7%, 93.8%, 98.9% and 93.7% of total residues in favoured region respectively in Ramachandran Plot (Figure 2B, 2D, 2F and 2H).

Geometrically acceptable values were obtained from modelled structures and most of residue torsion angels were present in allowed as well as favoured regions of plot. This indicated the thermodynamic stability of all the structures. Similarly, overall quality was identified using ERRAT It is also known as “overall quality factor” and high quality model range is > 50% [32]. The ERRATE scores for modelled alleles (DRB1*0701, DRB1*1101, DRB1*1301 and DRB1*1401) were 78.53, 77.24, 71.38, 77.84 (Figure 3). Thus, all results indicate the overall significant quality of all the modelled alleles.

---

Docking study

Docking study was performed between the antigen binding grooves of HLA alleles with identified T-cell epitopes. Docking scores of identified two promiscuous peptides were provided (Table 1). They showed average binding energies of -7 kcal/mol and had good binding scores. Molecular interactions of the peptides with predominant HLA-DRB1* alleles (receptor) among populations i.e. DRB1*0101 and DRB1*1501 were generated with LigPlot+ program. The identified peptides were having best docking scores and binding affinity with the HLA alleles of this study. They showed good number of hydrogen bond interactions. Therefore, predicted peptides have strong binding interactions with prevalent HLA molecules and interaction pattern of one peptide (S2) with respective alleles has been provided (Figure 4).

Figure 3A and 3B: ERRAT analysis plots of HLA alleles representing the overall qualities of modelled alleles:
A) DRB1*07:01 B) DRB1*11:01 C) DRB1*13:01 D) DRB1*14:01.

**Figure 4:** Docking result complexes between DRB1*0101 and DRB1*1501 and respective promiscuous peptide. A) S2, representing ligplot plus interactions between A and C-chains. B) Interactions between B and C-chains. C) Binding pattern of chain-C (peptide) in the groove of HLA-DRB1*0101 allele. D) S2, representing ligplot plus interactions between A and C-chains. E) Interactions between B and C-chains. F) Binding pattern of chain-C (peptide) in the groove of HLA allele.

<table>
<thead>
<tr>
<th>S. No</th>
<th>Code</th>
<th>Promiscuous T-cell epitopes</th>
<th>HLA-DRB1*receptors</th>
<th>Average score†</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>0101</td>
<td>0701</td>
</tr>
<tr>
<td>SMT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>S1</td>
<td>CRVLEFVRLAPGTY</td>
<td>-7.5</td>
<td>-7.7</td>
</tr>
<tr>
<td>2</td>
<td>S2</td>
<td>RVLEFVRLAPGTY</td>
<td>-8.1</td>
<td>-7.5</td>
</tr>
</tbody>
</table>

**Table 1:** Promiscuous T-cell epitopes identified from SMT protein of L. donovani and their docking results with all HLA alleles (kcal/mol).

**B-cell epitopes identification**

The B-cell epitopes were selected using the online servers including ABCpred, BCEPREDs, BepiPred and EBiPro. All the predicted epitopes were present on surface of protein. The epitopes present at least in three servers out of four, taken as promising B-cell epitopes and polar, non-polar, charged residues are shown (Table 2). The 3D structure of SMT protein was modelled using Swiss model [33]. The position of each predicted epitope was confirmed by visualizing on its 3D modelled protein (Figure 5).

<table>
<thead>
<tr>
<th>Sr. No</th>
<th>Position</th>
<th>Sequence</th>
<th>Molecular weight</th>
<th>PI</th>
<th>Linear Epitope</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>SMT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>4 - 19</td>
<td>ggRETapTNiRRRNK</td>
<td>1839.09</td>
<td>12.00</td>
<td>SB1</td>
</tr>
<tr>
<td>2</td>
<td>43 - 54</td>
<td>EERKaaTTTMvN</td>
<td>1350.51</td>
<td>6.24</td>
<td>SB2</td>
</tr>
<tr>
<td>3</td>
<td>111 - 119</td>
<td>CpyggpaRN</td>
<td>829.93</td>
<td>8.25</td>
<td>SB3</td>
</tr>
<tr>
<td>4</td>
<td>211 - 216</td>
<td>TDKYNp</td>
<td>736.78</td>
<td>5.50</td>
<td>SB4</td>
</tr>
<tr>
<td>5</td>
<td>231 - 239</td>
<td>gDjipEMET</td>
<td>948.01</td>
<td>3.57</td>
<td>SB5</td>
</tr>
<tr>
<td>6</td>
<td>266 - 281</td>
<td>SSpiKSpWYqphgD</td>
<td>1787.05</td>
<td>5.55</td>
<td>SB6</td>
</tr>
</tbody>
</table>

**Table 2:** Most promising linear B-cell epitopes of the SMT (SB) protein of L. donovani.

Hydrophilic residues: upper case letters, hydrophobic residues: lower case letters, charged residues: bold letters.

Discussion

The immune system of host organisms react to several foreign particles i.e. antigens and generate the efficient immune response against pathogens. Both T-cells and B-cells are involved in this process. This immune response can be replicated by synthetic peptides or epitopes against infectious organisms including Leishmania. The T-cell epitopes are recognized by host T-cells in association with specific HLA molecules and generate immune response [34]. In case of leishmaniasis, various promiscuous peptides have been designed using secretory or trans-membrane proteins of parasite [35]. Therefore, in the present study, we demonstrated that the immunogenic nature of strong vaccine candidate protein, SMT of L. donovani by using in-silico tools. In previous studies, it is demonstrated that the protection from the L. infantum by using SMT plus MPL®-SE vaccine that is generated Ag-specific Th1 immune responses [36]. SMT protein expressed in Leishmania parasite and absent in mammals, hence it is a potential vaccine candidate [4]. Antigenic features of protein was identified by secondary structures using SOPMA server and it was found that this protein possess high percentage of extended strands and random coils. There is advancement in research related to peptide-based vaccines [37]. It is studied that peptides able to bind to major histocompatibility complexes (MHC) with an affinity above a threshold value of 500 nm [38] act as T-cell epitopes. Hence, peptides selected on the basis of 100% binding affinity with all the HLA class-II alleles and which showed IC₅₀ values < 500 nM were selected as promiscuous peptides.

The 3D structures of DRB1*0701, DRB1*1101, DRB1*1301, DRB1*1401 were prepared by homology modeling using Modeller9V7. Modeller is most frequently used for homology or comparative protein structure modelling [26]. Homology modeling is considered as an important method of 3D structures generation and overall qualities can be studied using Ramachandran plot as well as ERRAT analysis (Kashyap, et al. 2016, 2017). Hence, overall qualities of modelled alleles were calculated using Ramachandran plot and ERRAT analysis and found to be good. Further, AutoDock vina was used for docking analysis of selected promiscuous peptides. Molecular docking is a key structure based, quick and accurate method for evaluating peptides binding to MHC molecules [39]. Docking analysis was done to explore the binding affinity of promiscuous epitopes/peptides. Promiscuous peptides of the study showed polar and non-polar residues and good solvent accessibility of the peptides. It was studied earlier that the diagnosis of VL is affected by cross reactions with other parasitic infections [40]. Therefore, the identification of Leishmania specific B-cell epitopes may be used for immunodiagnosis of Leishmania as well as anti-leishmanial vaccine development. We identified the linear B-cell epitopes using online server. Most promising B-cell epitopes were identified and their characteristic features were also calculated. Hence, the predicted T-cell and B-cell epitopes of the present study may be further used to explore for their ability to elicit protective immune response in-vitro/in-vivo.

Conclusion

In the present study, promiscuous T-cell and B-cell epitopes were predicted from SMT protein of L. donovani using in-silico tools. T-cell epitopes were identified as promising T-cell and B-cell epitopes which is able to bind HLA-class-II alleles, and have binding affinity to dif-

---

Epitope Prediction and Structural Analysis of Sterol 24-c-Methyltransferase Antigen of Leishmania donovani Using In Silico Approach

Different HLA molecules prevalent among the population. Whereas Linear B-cell epitopes may also be used for serodiagnostic purpose. This is the preliminary work and the father study is undertaken.

Disclosure Statement
The authors declare no conflict of interest.

Acknowledgements
The authors are thankful to the Shoolini University, Solan for providing the required facilities and resources to carry out this research work.

Bibliography

1. Desjeux P. "Leishmaniasis: current situation and new perspectives". Comparative Immunology, Microbiology and Infectious Diseases 27.5 (2004): 305-318.