



# Epitope prediction, modeling, and docking studies for H3L protein as an agent of smallpox

ELYAS MOHAMMADI<sup>1\*</sup>, SAMIRA DASHTY<sup>2</sup>

<sup>1</sup> Department of Animal Science, Faculty of Agriculture, Ferdowsi University of Mashhad, Mashhad, Iran

<sup>2</sup> Department of Medical Science, University of Medical Sciences of Mashhad, Mashhad, Iran

## Abstract

Despite the eradication of *Variola* and *Vaccinia viruses*, smallpox is still a potential threat to human societies. The H3L protein is conserved and immune-dominant among different strains of poxviruses. The aim of this study was to detect epitope regions in the H3L protein by bioinformatics tools and examine the accuracy of an experimentally determined epitope, VP35#1, against HLA-A\*0201 by molecular docking. H3L epitopes were predicted against major histocompatibility complex (MHC) II receptors by bioinformatics servers. Antigenic potency of these epitopes was investigated using Vaxijen 2.0 software. The digestibility of predicted epitopes was predicted using online servers. Epitopes that were missed in the partial crystallographic structure of H3L protein were modeled using Muster server. It has been determined that H3L protein, including 8 predicted epitopes for 6 MHC Class II receptors, is conserved among the agents of smallpox. The antigenicity of the epitopes has been confirmed through Vaxijen 2.0 software. The binding affinity for docking VP35#1 and the HLA-A\*0201 receptor ( $\Delta G = -7.4$ ) showed stable and powerful interaction.

**Key words:** bioinformatics, receptor, MHC II, epitope, smallpox, H3L protein

## Introduction

In the last 30 years, smallpox, an infectious disease with a high mortality rate caused by dsDNA viruses belonging to Orthopoxs genus, has been mostly eradicated; however, it still constitutes a threat to human population. This threat is not only due to the possibility of a deliberate outbreak and bio-terrorism but also due to the possibility of a natural outbreak of this disease (Kennedy et al., 2009). *Variola virus* and *Vaccinia virus* (VACV) are the main causes of smallpox (Fenner, 1988). H3L is a 35-kDa VACV envelope protein encoded by open reading frame (ORF) H3L (Drexler et al., 2003). This protein is highly immuno-dominant and conserved among the agents of smallpox (Enserink et al., 2002). The deletion of the H3L ORF interferes with virus maturation and reduces virus infectivity and virulence; therefore, the H3L protein may have roles in both virus maturation and entry (Lin et al., 2000).

In the 1700s, Edward Jenner proved that infection with the *Cowpoxvirus* can protect humans against smallpox and he prepared the stage for vaccination in general (Jenner's work represented the first scientific attempt to control an infectious disease by the deliberate use of vaccination). However, he did not discover vaccination; in fact, he was the first person to confer a scientific status on the procedure and to pursue its scientific investigation (Riedel, 2005). Nowadays, vaccination is a common practice, and recently all vaccines for smallpox have been made from one type of Orthopox virus called VACV (Kennedy et al., 2009). As an example, the *Modified Vaccinia Virus Ankara* is a highly attenuated candidate for smallpox vaccine which is derived from VACV (Drexler et al., 2003). Vaccination with VACV provides long-term protection against smallpox (Terajima et al., 2003).

An epitope is believed to be a part of an antigen that causes an immune response (Huang and Honda, 2011).

\* Corresponding author: Department of Animal Science, Faculty of Agriculture, Ferdowsi University of Mashhad, Mashhad, Iran; e-mail: e.mohammadi@mail.um.ac.ir

Epitopes are short peptide sequences divided into two general categories according to their specific receptor: 1) identifiable by humoral immunity with the help of lymphocytes B or 2) detectable by cellular immunity with the help of lethal T cells ( $CD8^+$  T), and helper T cells ( $CD4^+$  T) (Chen et al., 2011). In smallpox cellular immunity, helper T cells ( $CD4^+$  T) are particularly important.  $CD4^+$  T cells protect humans against smallpox, and  $CD8^+$  T cells produce a strong memory for this disease (Drexler et al., 2003). Based on the studies conducted on major histocompatibility complex (MHC) class I receptors and the HLA-A\*0201 allele, an epitope that is conserved among all poxviruses has been found (Enserink et al., 2002; Kennedy et al., 2007; Tang et al., 2008). The "SLSAYIIRV" sequence (VP35#1), as an epitope in the H3L protein, is specific to HLA-A\*0201, an allele of MHC I (Enserink et al., 2002; Kennedy et al., 2007; Tang et al., 2008). Meanwhile, with the presence of epitopes of the H3L protein specific to MHC class II receptors, the combination of these two categories of epitopes can lead to the production of a peptide vaccine against smallpox (Terajima et al., 2003). Due to the risks of using conventional vaccines for smallpox, advances in recombinant vaccines and improving their effectiveness seem necessary (Enserink et al., 2002). Some of the disadvantages of conventional vaccines against smallpox are: 1) the need for a cold chain (the need to continually keep vaccines at temperatures between 2–8°C) to maintain their potency, 2) the need for hiring skilled healthcare workers, which increases labor costs to produce conventional vaccines, 3) the risk of infection caused by the usage of live vaccine (Giddens et al., 1971).

The main objective of the present study was to investigate the H3L protein (a smallpox agent) for the prediction of epitopes using bioinformatics analyses and to identify the epitopic regions specific to MHC class II receptors. Such epitopes can be used to produce recombinant vaccines instead of using whole antigens for vaccination. The other objective of the bioinformatics survey was to verify the validity of the experimental epitope VP35#1 for the HLA-A\*0201 receptor via docking studies (Enserink et al., 2002; Kennedy et al., 2007; Tang et al., 2008).

## Material and methods

### *Epitope prediction of the H3L protein*

Considering the importance of  $CD4^+$  T cells in creating resistance to smallpox disease (Drexler et al.,

2003), a sequence of the H3L protein as immuno-dominant and conserved among the agents of smallpox was extracted from the National Center for Biotechnology Information (NCBI) database (AIX99220.1 Accession number).

To predict the specific epitopes of the H3L protein toward MHC class II receptors (Fig. 1), three softwares (using three different algorithms) listed in Table 1 were used. First, 30 epitopes with the highest ranks predicted by each server were selected. Then, only the epitopes with the highest scores that were common for prediction by all the three servers were chosen. The method of the immune epitope database (IEDB) server for epitope prediction was IEDB. IEDB recommends the Consensus method consisting of NetMHC, stabilized matrix method (SMM) and scoring matrices derived from combinatorial peptide libraries (CombLib). The algorithms of these softwares are based on the prognosis of the T-cell epitopes. In these servers, the amino acid sequences of the selected protein are used to determine regions with a high probability of epitope properties. Human leukocyte antigen DRB (HLA-DRB) alleles (0101, 0301, 0401, 0701, 1101, and 1501) were selected for epitope prediction based on their frequency and their impact on various diseases (Collins, 2000; Mack et al., 2009; Rossman et al., 2003).

### *Modeling the predicted epitopes of the H3L protein from the primary to secondary structure*

The crystallographic structure of the H3L protein was downloaded from The Research Collaboratory for Structural Bioinformatics (RCSB) database (5EJ0) in pdb format. Considering that this crystallographic structure of the H3L protein is partial, the predicted epitopes numbered 2, 3, 5, and 6, shown in Table 2, were present in the crystallographic structure, whereas epitopes numbered 1, 4, 7, and 8 were absent from the crystallographic structure of the H3L protein. The three-dimensional structures of the absent epitopes were modeled using the Muster (Wu et al., 2008) server (a powerful protein modeling server) under the *ab initio* procedure.

### *Determining the antigenicity of the predicted epitopes of the H3L protein*

The final selection of H3L predicted epitopes was performed using Vaxijen 2.0 (Doytchinova and Flower 2007a, 2007b, 2008) software with a 0.6 threshold. This software classified epitopes based on their physicochemi-

**Table 1.** Servers used to predict MHC class II epitopes

Server name	URL	Technique/Algorithm	Class	References
SYFPEITHI	http://www.syfpeithi.de/Scripts/MHCServer.dll/EpitopePrediction.htm	Motif Matrices (MM)	I and II	Rammensee et al., 1999
PROPPRED	http://www.imtech.res.in/raghava/propred/	QM (Virtual QM)	II	Singh et al., 2001
IEDB	http://tools.iedb.org/mhcii/	ANN-regression/SMM-QM	II	Zhang et al., 2008

```
>Vaccinia virus IMV heparin binding surface protein [AIX99220.1]
MAAVKTPVIVVPVIDRPPSETFPNVHEHINDQKFDVVDKNEVMPEKRNVVVVEDDPDHKYDAFIQWTGG
NIRNDDKYTHFFSGFCNTMCTEETKRNIARHLALWDSNFFTELENKKVEYVVIVENDNVIEDITFLRPVL
KAMHDKKIDILQMRIITGNKVKTELVMKDHAIFTYTGGYDVLSAYIIRVTTALNIVDEIIKSGGLSS
GFYFEIARIENEMKINRQILDNAAKYVEHDPRLVAEHRFENMKPNFWSRIGTAAAKRYPGVMYFTFTPLI
SFFGLFDINVIGLIVILFIMFMLIFNVKSKLLWFLTGTFVTAFI
```

**Fig. 1.** Amino acid sequence of H3L protein**Table 2.** H3L predicted epitopes against MHC class II receptors

Number	Epitope	Vaxijen score	Receptor	Modeling	Start	End
1	FMLIFNVKSKLL	1.1382	HLADRB1-0101	Muster (AB ANITIO)	301	312
2	LSAYIIRVTTALNI	0.6355	HLADRB1-0101	X-ray Diffraction	185	199
3	GGNIRNDDKYTHF	1.0661	HLADRB1-0301	X-ray Diffraction	69	81
4	FIMFMLIFNVKSKL	1.1187	HLADRB1-0401	Muster (AB ANITIO)	298	311
5	LSAYIIRVTTALNI	0.6355	HLADRB1-0401	X-ray Diffraction	185	199
6	LSAYIIRVTTALNI	0.6355	HLADRB1-0701	X-ray Diffraction	185	199
7	FIMFMLIFNVK	1.2092	HLADRB1-1101	Muster (AB ANITIO)	298	308
8	IVILFIMFMLIFNVK	0.8438	HLADRB1-1501	Muster (AB ANITIO)	294	308

cal properties. The peptides with higher scores compared to the threshold were identified as epitopes. The scores given by the server to the epitopes are shown in Table 2.

#### **Predicting the enzymatic digestion of the predicted epitopes of the H3L protein**

The possibility of the predicted candidate epitopes being digested was tested using <http://db.systemsbio.net/server> by Proteomics Toolkit software. In this server, trypsin, chymotrypsin, elastase, clostripain, cyanogen bromide, iodosobenzoate, proline endopept, staph protease, and aspartate-n-terminal (*AspN*) enzymes were used to determine the digestibility of the epitopes.

#### **Docking studies of VP35#1 epitope and HLA-A\*0201 allele**

3UTQ complex composed of HLA-A\*0201 allele (the chosen allele) and an epitope (chosen for the study) was

retrieved from the RCSB database. This complex for docking studies was modified using Pymol 1.8 (The PyMOL Molecular Graphics System, Version 1.8, Schrödinger, LLC.) (DeLano, 2002) and Chimera 1.11.2 (Pettersen et al., 2004). For this purpose, the water and ligand binding molecules (undesired epitope) were removed from the receptor by the Pymol software, polar hydrogen was then added to the receptor, and the energy was then optimized by the Chimera software. The crystallographic structure of H3L protein was downloaded from the RCSB database (5EJ0), and the structure of VP35#1 epitope was extracted using the Pymol software and stored in pdb format. The preparation of VP35#1 epitope for docking was performed in a manner similar to that of the 3UTQ complex (removing water, adding polar hydrogen, and finally energy optimization). The docking study of the VP35#1 epitope and the HLA-A\*0201 receptor was

performed using the Pyrx software (Autodock-vina) (Trott et al., 2010). This process was repeated 3 times, and the average of binding affinity ( $\Delta G$ ) was calculated and reported in Table 4. All amino acids involved in binding the epitope-receptor and the length of the bonds were reported for all three replicates in Table 4.

#### **Validation of our docking procedure and epitope prediction servers**

3UTQ is a crystallographic structure of a complex containing HLA-A\*0201 receptor and ALWGPDPAAA epitope (this is not the epitope chosen for docking studies in this research, but used for docking validation) which is effective in type 1 diabetes (Bulek et al., 2012). This structure was obtained by the X-ray diffraction method and it shows the correct position for ligand binding in HLA-A\*0201. Considering the fact that “ALWGPDPAAA” is an experimentally confirmed epitope specific to HLA-A\*0201 allele (Bulek et al., 2012), it is expected that docking the components of 3UTQ (epitope and receptor) will show a considerable interaction (Osman et al., 2016). To validate the docking method and software, we separated the components of this complex (3UTQ) using the Pymol software. Then, as mentioned above, with the help of Chimera software we prepared the components for docking, and then docked them (HLA-A\*0201 receptor and ALWGPDPAAA epitope) with Pyrx software under the Autodock-Vina procedure with three replications.

We verified the validity of the epitope prediction software used in this research with the following genetically modified antigens 1) GroEL (belonging to the chaperonin family of molecular chaperones and found in a large number of bacteria), 2) 70 kDa heat shock proteins (Hsp70s or DnaK) (family of conserved, ubiquitously expressed heat shock proteins), 3) Omp31 (a 31 kDa outer-membrane immunogenic protein from *Brucella melitensis*) and 4) superoxide dismutase antigen of *Brucella abortus*, which were previously confirmed in experimental studies (Forouharmehr et al., 2015). In this case, we tried to determine if the epitope prediction softwares used in this research were capable of predicting the genetically modified epitopes of these antigens or not.

## **Results**

### **Investigating the conservation of H3L protein**

The H3L protein sequences from the *Cow poxvirus* (ADZ29223.1), VACV (AIX99220.1), *Variola virus* (AAA

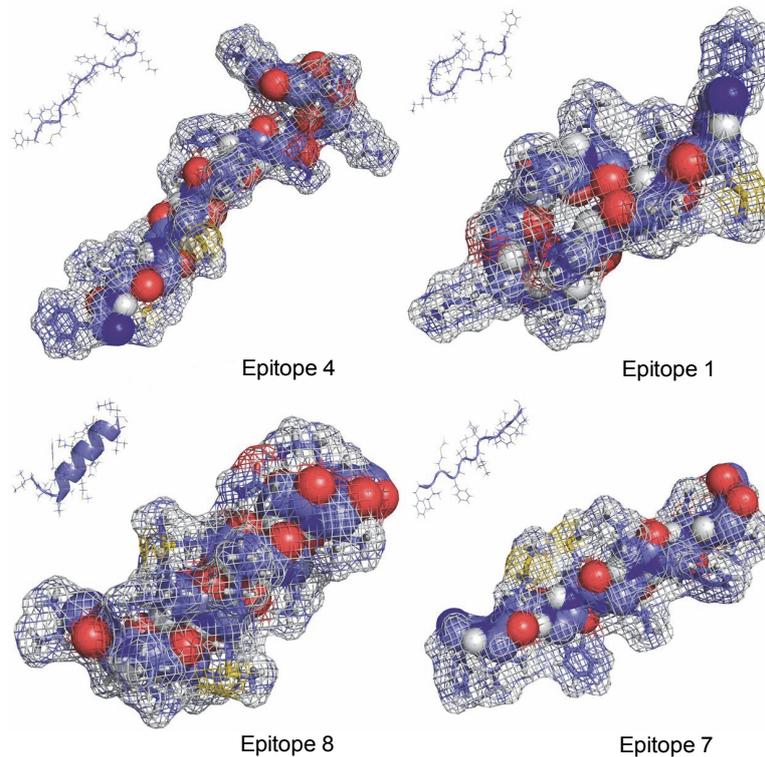
60834.1), and *Monkey poxvirus* (NP\_536520.1) strains were extracted from the NCBI database (<https://www.ncbi.nlm.nih.gov/>) and compared (shown in Fig. 2) to investigate their conservation. In this case, 24 single-nucleotide differences were observed along the 324 amino acid alignments of these four sequences. In addition, one gap was observed due to an extra Asparagine in *Variola virus* sequence. On comparing *Variola virus*, *Monkey poxvirus*, and *Cow poxvirus* to VACV, 24 single-nucleotide differences in the H3L protein sequence (16, 5, and 3 single-nucleotide differences respectively) were identified in just 1 (*Variola virus*), 2 (*Variola virus*, *Monkey poxvirus*), and all three sequences (*Variola virus*, *Monkey poxvirus*, and *Cow poxvirus*) as compared to the VACV. Epitope mapping along the VACV sequence indicated that all predicted epitopes were in completely conserved parts of the protein. Just one single-nucleotide difference between VACV and *Monkey poxvirus* was observed in epitope 3 (isoleucine from VACV was changed to aspartic acid in *Monkey poxvirus*).

### **The prediction of epitopes of the H3L protein against a number of HLA-DRB alleles**

HLA-DRB alleles (0101, 0301, 0401, 0701, 1101, and 1501) were used to identify the epitopes of the H3L protein specific to the MHC II receptor based on their frequency between different populations and their importance in other diseases (Table 1). Generally, 8 epitopes have been predicted in the H3L protein for 6 MHC class II variants. Epitope number 2, 5, and 6 are the same but they were predicted 3 times for three different variants of MHC II receptors. Thus, epitopes 1 and 2 were specific to HLA-DRB1-0101, epitope 3 was specific to HLA-DRB1-0301, epitopes 4 and 5 were specific to HLA-DRB1-0401, and epitopes numbered 6, 7, and 8 were specific to HLA-DRB1-0701, HLA-DRB1-1101, and HLA-DRB1-1501, respectively (Table 2).

Epitopes predicted for HLA-DRB variants (0101, 0301, 0401, 0701, 1101, 1501) were confirmed by VAXIJEN 2.0 software, with a 0.6 threshold. Through Vaxijen software, epitopes 1, 3, 4, 7, and 8 (Table 2) were identified as antigens with a threshold higher than 1. Epitope 8 and also epitopes 2, 5, and 6 (Table 2), which are the same, were identified as antigens with a threshold higher than 0.8 and 0.6, respectively. The 8 epitopes given in Table 2 were selected and used for the prediction of enzymatic digestion, and structure modeling.





**Fig. 3.** Modeling of predicted epitopes that were missed in the crystallography of H3L protein

#### ***Prediction of enzymatic digestion of selected epitopes***

For all epitopes presented in Table 3, computer simulation predicted the possibility of digestion by chymotrypsin and Elastase enzymes using <http://db.systemsbiology.net/> server through Proteomics Toolkit software. Epitopes other than 7 and 8 were digested by trypsin, and epitopes number 1, 4, 7, and 8 were digested by cyanogen\_bromide. The only epitope that could be digested by *AspN* was epitope number 4. Eventually, the epitopes numbered 2, 5, and 6 were digested by clostripain.

#### ***Modeling of the predicted epitopes from the primary to secondary structure***

As the importance of peptides and their function can be associated with their secondary and tertiary structures, we used the Muster server (Wu et al., 2008) under *ab initio* procedure for modeling of the predicted epitopes. The crystallographic structure of H3L (5EJ0) downloaded from the RCSB database was made for an incomplete protein. Due to the lack of epitope structures numbered 1, 4, 7, and 8 in 5EJ0, we modeled these epitopes using the Muster server (Fig. 3).

The modeled epitopes (no. 1, 4, 7, and 8) and the spaces occupied by their side chains are shown in Fi-

gure 3. In these epitopes, the carbon atom is presented in pale blue, hydrogen in gray, nitrogen in blue, oxygen in red, and sulfur in yellow.

Epitope 1 (“FMLIFNVKSKLL”) was predicted to possess a random coil structure. The linear structure of this epitope was continued till the 7th amino acid, valine, and then it was interrupted by the presence of the lysine amino acid after valine. The structure of epitope 4 (“FIMFMLIFNVKSKL”) was predicted as linear from the beginning till its 11th amino acid, lysine. Due to the interaction between the side chains of the last 4 amino acids of this epitope, the linear structure was disturbed. Epitope 7 (“FIMFMLIFNVK”) had a linear structure and the structure predicted for epitope 8 (“IVILFIMFMLIFNVK”) differed from the structures of other epitopes. Due to its longer length (15 amino acids as compared to 12, 14, and 11 amino acids in epitopes 1, 4, and 7, respectively) in comparison to other modeled epitopes, the structure was predicted to be an  $\alpha$ -helix.

#### ***Validation of docking experiments***

The average of  $\Delta G$  for replications of docking of HLA-A\*0201 as a receptor and “ALWGPDPAAA” as a ligand using Pyrx software was determined to be

**Table 4.** Three replications (R) and Analysis for docking of VP35#1 Epitope and HLA-A\*0201 receptor

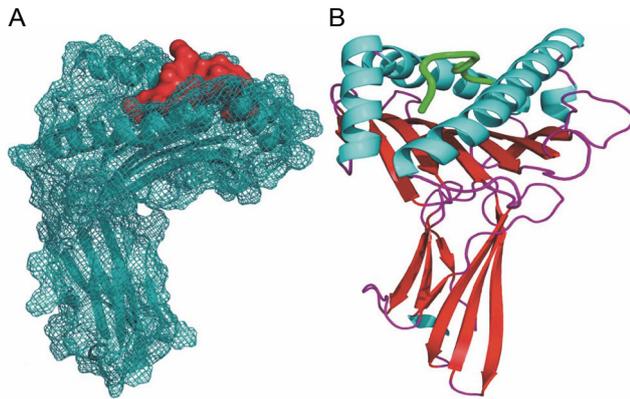
Epitope			Receptor		
R1	R2	R3	R1	R2	R3
S (1)	S (1)	S (1)	T (143)	K (66)	H (70)
L (2)	L (2)	L (2)	Y (84)	D (77)	Y (171)
S (3)	R (8)	I (6)	W (147)	Y (99)	Y (159)
R (8)	V (9)	R (8)	D (77)	W (147)	Y (7)
			T (73)	Q (155)	T (73)
			Y (159)		D (77)
			E (63)		
Bond length [Å]					
R1		R2		R3	
T (143): S (1): 2.5		Q (155): R (8): 2.2		Y (99): V (9): 2.1	
Y (84): S (1): 2.4		K (66): R (8): 2.5		Y (171): S (1): 2.1	
W (147): S (1): 2.1		W (147): S (1): 2.4		Y (159): S (1): 2.5	
D (77): L (2): 2.1		W (147): L (2): 2.3		H (70): L (2): 2.6	
T (73): S (3): 2.3		D (77): L (2): 2.2		Y (7): S (1): 2	
Y (159): R (8): 1.9		D (77): S (1): 2.1		T (73): I (6): 2.2	
E (63): R (8): 2.3				D (77): V (9): 2.2	
E (63): R (8): 2.3					
Molecular docking repeats			The strength of docking for epitope SLSAYIIRV and the receptor HLA-A 0201 ( $\Delta G$ )		
1			-7.2		
2			-7.3		
3			-7.5		
Average			-7.33		
Validation repeats			The strength of docking for epitope ALWGPDPA AAA and the receptor HLA-A 0201 ( $\Delta G$ )		
1			-9.6		
2			-11.5		
3			-10.2		
Average			-10.43		

-10.43 KJ/mol. These values of docking are considered as evidence to prove that our docking software worked properly, because "ALWGPDPA AAA" is an experimentally confirmed epitope specific to HLA-A\*0201 receptor. So, it can be concluded that the software we used for docking in this study could mimic a real phenomenon (interaction between "ALWGPDPA AAA" and HLA-A\*0201). The docking results of this validation are shown in Table 4. To validate all epitope prediction, servers used in this study were verified on previously identified relevant epitopes of GroEL, Dnak, Omp31, and SOD antigens. Also, SYFFPEITHI server ([www.syfpeithi.de](http://www.syfpeithi.de), proprietary for both MHC class I and II receptors) could predict the experimentally confirmed epitope VP35#1 (Enserink et al., 2002; Kennedy et al., 2007; Tang et al., 2008) in the H3L protein, thus confirming the accuracy of this epitope prediction software.

#### **Molecular docking of VP35#1 and HLA-A\*0201**

The docking operation was performed with the use of Pyrx software under the Autodock-vina procedure for the VP35#1 epitope and the HLA-A\*0201 receptor in three replications. The binding affinities for three replications were -7.4, -7.3, and -7.5 KJ/mol. The average of binding affinities for the three repeats (-7.33 KJ/mol) and placing this epitope in the correct position of the receptor (between two  $\alpha$ -helix chains, see Fig. 4) revealed that VP35#1 had a considerable interaction and the correct position in HLA-A\*0201.

The three different positions of epitopes in the docking groove of the receptor during docking studies are shown in Figure 5. In all three replications of docking, Asp-77 from the  $\alpha$ -helix chain of the receptor was involved in interactions with different residues of the epitope,



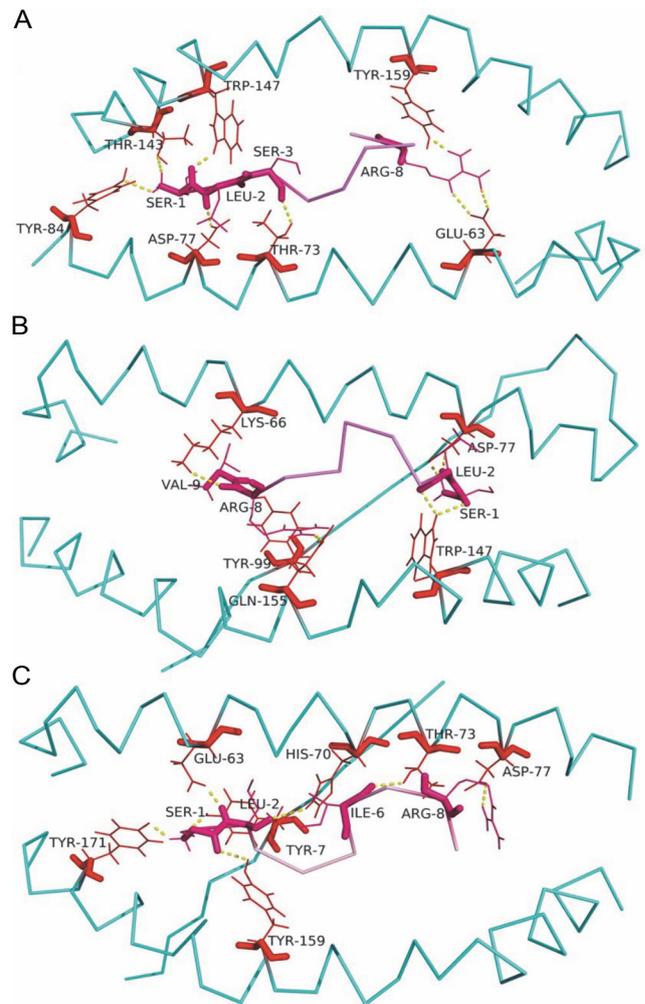
**Fig. 4.** The figure of the complex for VP35#1(epitope) and HLA-A\*0201(receptor). In Part A, the secondary structure of the protein is characterized by a distinct color. The  $\alpha$ -helix chains are marked in pale blue,  $\beta$  sheets reddish, and loop structures in violet color. The epitope is also visible in green. In part B, the receptor and the areas occupied by its side chains are marked in pale blue and the epitope marked in red

which shows the importance of this residue in the docking process.

The amino acids that were involved in interactions between VP35#1 and HLA-A\*0201 are listed in Table 4.

## Discussion

Using vaccination as a vital tool for preventing a wide spread of diseases has resulted in substantial measures taken to combat many infectious diseases such as influenza, smallpox, varicella, pertussis, diphtheria, tetanus, polio, hepatitis, and rotavirus. The term “epitope-driven vaccine” is an attractive concept that is being successfully followed in a large number of research groups (Soria-Guerra et al., 2015). Foo et al. (2008) used POPRED software (Singh et al., 2001), one of the epitope prediction softwares that we used in our study, to predict the epitopic regions of the major antigen (p1) of enterovirus-71. They could predict 3 epitopes specific to 25 HLA-DR alleles. The synthetic peptides of these epitopes were tested for their capacity to induce  $CD4^+$ T proliferation in five human volunteers.  $CD4^+$ T-cell proliferative responses were observed in all of them. It evidenced from POPRED software efficiency studies that the epitopes predicted through this software were efficient in experimental studies (Foo et al., 2008). Molero-Abraham et al. (2015) used position-specific scoring matrices to predict common epitopes among agents of smallpox. The number of predicted epitopes for MHC II



**Fig. 5.** Three replications of docking between VP35#1 (epitope) and  $\alpha$ -helix chains of HLA-A\*0201 (receptor).  $\alpha$ -helix chains are colored in light blue. The residues of the receptor that had an interaction with the epitope are colored in red and the residues of the epitope that had an interaction with the receptor are colored in bold pink. The main chain of residues that had an interaction is shown as sticks and the side chains are shown as lines. The other residues are shown as ribbons

receptors was much smaller than that for MHC I receptors. For epitope prediction, the authors used H3L protein, but none of the epitopes predicted in our study coincided with the ones described by Molero-Abraham et al. (2015). On the other hand, the authors predicted the VP35#1 epitope for HLA-A\* 0201. Drexler et al. (2003) selected HLA-A\*0201-restricted epitopes derived from VACV proteins (one of these proteins was H3L protein) using the SYFPEITHI epitope prediction server (Rammensee et al., 1999), one of the epitope prediction softwares we used in our study. They predicted the VP35#1 epitope and evaluated its ability to induce cel-

lular immunity, especially CD8<sup>+</sup>T cells. It can be concluded that a number of computational and experimental studies considered VP35#1 as an epitope specific to HLA-A\* 0201 allele which can stimulate CD8<sup>+</sup>T cells.

According to a study by Drexler et al. (2003), the H3L protein is highly conserved among agents of smallpox, including *Variola* and *Vaccinia*, and stimulates the immune system; therefore, H3L was used to screen for the presence of specific epitopes for MHC class I receptor. The VP35#1 epitope, dedicated to the HLA-A\*0201 allele, does not provide protection against smallpox despite the creation of a wide range of modified CD8<sup>+</sup>T cells (Enserink et al., 2002). The selection of the H3L protein as a candidate for epitope prediction was done based on previous studies indicating that this 35-kDa VACV envelope protein encoded by ORF H3L is immuno-dominant and conserved among poxviruses (Da Fonseca et al., 2000; Zinoviev et al., 1994). VP35#1 epitope is a part of the H3L protein and its conservation among the known H3L gene homologues derived from various poxviruses was investigated by Lee et al. (2001). This investigation revealed a complete conservation of the "SLSAYIIRV" sequence in VP35#1 among Orthopoxvirus species. Drexler et al. (2002) showed that this epitope is specific for HLA-A\* 0201. In the present study, we validated and confirmed this result (specificity of VP35#1 to HLA-A\* 0201) by a bioinformatics procedure. On the other hand, we showed that "LSAYIIRVTTALNI" peptide has epitope property for HLA-DRB1 0101, 0401, and 0701 variants. This epitope is very similar to VP35#1, which is specific to HLA-A\* 0201, but VP35#1 individually did not show epitopic properties for HLA-DRB alleles.

To identify the epitopes of VACV antigens, it is first necessary to know which of the B cells (that are involved in humoral immunity), lethal T-cell (CD8<sup>+</sup>T) or T-helper cells (CD4<sup>+</sup>T) (that play a role in cellular immunity) have the most important functions in the process of controlling the pathogenicity of this virus by the immune system. Severe and life threatening complications caused by vaccination with VACV were associated with congenital or acquired T-cell deficiencies, but not with congenital agammaglobulinemia which results in a lack of gamma globulin in the blood plasma (Terajima et al., 2003). Historically, individuals with a defective humoral immunity who received a smallpox vaccine developed normal characteristics after vaccination, but were at a higher

risk of developing complications such as generalized vaccinia (a cutaneous condition that occurs 6–9 days after vaccination, characterized by a generalized eruption of skin lesions and caused by the VACV). In contrast, individuals with T-cell defects failed to contain the infection after immunization and were at the risk of serious complications, including vaccinia necrosum (painless but progressive necrosis and ulceration, Kennedy et al., 2009). In another study, vaccination against smallpox in HLA transgenic mice (VennVax) resulted in the stimulation of considerable T-cell responses (Moise et al., 2011). No antibody response pre-challenge was observed, either against the whole vaccinia antigens or vaccine epitope peptides. Remarkably, 100% of the vaccinated mice survived the lethal vaccinia challenge, demonstrating that protective immunity to vaccinia does not require B cell priming (Moise et al., 2011).

In mouse system studies, mice with their B cells suppressed as well as mice with suppressed  $\beta$ 2-microglobulin (unable to produce MHC class I receptors) were immunized as a result of vaccination (Wyatt et al., 2004). On the other hand, mice with reduced (via mutation) MHC class II receptor levels and double knocking mice with decreased levels of MHC class II and MHC class I receptors were poorly immunized or were not immunized by vaccination at all (Wyatt et al., 2004). It can be concluded that B cells and CD8<sup>+</sup>T cells are less important than CD4<sup>+</sup>T cells in developing an immune response against the smallpox virus. Therefore, the importance of T cells to confront smallpox, especially CD4<sup>+</sup>T cells, is higher than that of B cells, and when the vaccination is done in the absence of T cells in the body, the immune response is not well-established (Drexler et al., 2003; Enserink et al., 2002).

Severe complications in people, caused by vaccination with congenital T-cell deficiency, indicate the importance of T cells in the prevention of smallpox. Among T cells, the importance of CD4<sup>+</sup>T cells is greater than that of CD8<sup>+</sup>T cells. However, the role of CD8<sup>+</sup>T cells cannot be ignored, because the memory formed by CD8<sup>+</sup>T cells against smallpox is more stable than the memory created by CD4<sup>+</sup>T cells (Drexler et al., 2003; Enserink et al., 2002; Kennedy et al., 2007; Tang et al., 2008).

Selected alleles for epitope prediction had either high frequency among MHC II alleles in different human populations or played important roles in the case of antigen presentation and causing immunity against other disea-

ses (Collins, 2000; Mack et al., 2009; Rossman et al., 2003). HLA-DRB1 0101 is the most common allele of MHC II receptors in the Caucasian, Asian/Pacific Islander, and Native American groups (Collins, 2000). Based on a study on American human population conducted by Rossman et al. (2003), the frequencies of HLA-DRB alleles (0101, 0301, 0401, 0701, 1101, and 1501) were 0.0527, 0.1002, 0.0675, 0.1086, 0.0686, and 0.0770, respectively. On the other hand, HLA-DRB1 1101 and 1501 were identified as risk factors for sarcoidosis disease (Rossman et al., 2003).

As a result of our study, we predicted the existence of 6 epitopes in the H3L protein (conserved and immuno-dominant envelope protein that binds to the cell surface and causes the attachment of the smallpox virion to the target) for MHC II receptors of antigen-presenting cells. Due to the high frequency of occurrence of HLA-A\* 0201 allele in populations (0.28) (Mack et al., 2009), the importance of using this epitope for evoking CD<sup>+</sup>8 T cells is undeniable. On the other hand, considering the 0.0527, 0.1002, 0.0675, 0.1086, 0.0686, and 0.0770 frequency of occurrence of HLA-DRB1 0101, 0301, 0401, 0701 1101, and 1501 variants of MHC II receptors, respectively, in different human populations, (Collins, 2000; Mack et al., 2009; Rossman et al., 2003) the usage of epitopes of H3L protein which are specific to MHC II receptors can evoke CD<sup>+</sup>4 T-cell response. Overall, using both of the VP35#1 (specific to MHC I, HLA-A\* 0201 allele) and epitopes we predicted for MHC II receptors can be the basis for the production of a high performance recombinant vaccine against smallpox.

In the tertiary structure of MHC receptors, there is a groove between 2  $\alpha_1$  and  $\alpha_2$  domains with an  $\alpha$ -helix structure  $\sim 25\text{\AA}$  long and  $\sim 10\text{\AA}$  wide which constitutes a specific interaction site for epitopes (Hewitt, 2003). There are several pockets in this binding groove, in which case their role is to stabilize the epitope-receptor interaction by burying some amino acid anchors of the epitopes in themselves. These pockets play their role by interaction of their residue side chains with the anchor of the epitopes. According to the results of our docking experiments, VP35#1, an epitope of the H3L protein, could be attached in the HLA-A\*0201 binding groove (Fig. 5). Importantly, only if the side chains of the epitopes are placed in the appropriate receptor pocket and establish a strong interaction with it, the receptor would be able to present the epitope to the cellular immunity

components, and therefore stimulate the immune system.

Based on our results of modeling the predicted epitopes in H3L protein, most of the models had a linear structure (Fig. 3) and these results are in agreement with observations that specific epitopes for cellular immune receptors are more linear (Patronov et al., 2013). The specific epitopes for the MHC II receptor are usually between 12 and 15 amino acids long and the specific epitopes of the MHC I receptor are 8 to 11 amino acids long (Patronov et al., 2013). The lengths of epitopes 1, 4, 7, 8, which were specific to MHC II receptors, were 12, 14, 11, and 15 amino acids, respectively, while the length of VP35#1 was 9 amino acids. It confirms the report of Patronov et al. (2013) that the length of epitopes for MHC II is longer than that of the epitopes specific to MHC I receptors.

## Conclusions

The epitope prediction accuracy of an antigen might be increased by using three different servers with different algorithms (like using PROPPRED, IEDB, and SYFFPEITHI servers in our study) as compared to using a single server and algorithm for epitope prediction. In this case, the epitopes that were predicted by all three servers with top scores were chosen. This procedure might have increased the probability of epitopic properties of a predicted epitope for experimental studies. Sifting predicted epitopes with the Vaxijen server might increase the accuracy of the predicted epitopes.

Considering the high frequency of HLA-A\*0201 allele (0.28) among MHC I alleles in different human populations (Mack et al., 2009) and the high frequency (the frequency reported in the discussion) of MHC II alleles that we used for epitope prediction of the H3L protein in our study (Collins, 2000; Mack et al., 2009; Rossman et al., 2003), it can be concluded that using predicted epitopes (8 predicted epitopes of H3L protein for their specific MHC II receptors) and VP35#1 (specific to HLA-A\*0201 as a MHC I allele) as a polytope may generate a significant immune response by evoking cellular immunity, and it may immunize the body against the smallpox disease as such.

Based on the average of binding affinity ( $\Delta G = -7.33$  KJ/mol) for three replications of the docking between the VP35#1 epitope and the HLA-A\*0201 allele of MHC I, it can be concluded that in experimental stu-

dies the epitope may have a powerful interaction with the binding groove of HLA-A\*0201 allele, and that the epitope-receptor complex can be introduced to CD8<sup>+</sup> T cells, which cause immune response.

## References

- Bulek A.M., Cole D.K., Skowera A., Dolton G., Gras S., Madura F., Drijfhout J.W. (2012) *Structural basis for the killing of human beta cells by CD8<sup>+</sup> T cells in type 1 diabetes*. *Nat. Immunol.* 13(3): 283–289.
- Collins M., Tang T., Slack R., Sintasath D., Hartzman R., Ng J., Hurley C. (2000) *The relative frequencies of HLA DRB1\*01 alleles in the major US populations*. *HLA* 55: 48–52.
- Chen P., Rayner S., Hu K.H. (2011) *Advances of bioinformatics tools applied in virus epitopes prediction*. *Virol. Sin.* 26(1): 1–7.
- Da Fonseca F.G., Wolffe E.J., Weisberg A., Moss B. (2000) *Characterization of the vaccinia virus H3L envelope protein: topology and posttranslational membrane insertion via the C-terminal hydrophobic tail*. *J. Virol.* 74(16): 7508–7517.
- DeLano W.L. (2002) *Pymol: An open-source molecular graphics tool*. *CCP4 Newslett. Prot. Crystal.* 40: 82–92.
- Drexler I., Staib C., Kastenmüller W., Stevanović S., Schmidt B., Lemonnier F.A., Sutter G. (2003) *Identification of vaccinia virus epitope-specific HLA-A\*0201-restricted T cells and comparative analysis of Smallpox vaccines*. *Proc. Natl. Acad. Sci. USA* 100(1): 217–222.
- Doytchinova I.A., Flower D.R. (2007a) *VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines*. *BMC Bioinf.* 8(1): 4.
- Doytchinova I.A., Flower D.R. (2007b) *Identifying candidate subunit vaccines using an alignment-independent method based on principal amino acid properties*. *Vaccine* 25(5): 856–866.
- Doytchinova I.A., Flower D.R. (2008) *Bioinformatic approach for identifying parasite and fungal candidate subunit vaccines*. *Open Vaccine J.* 1(1): 4.
- Enserink M. (2002) *In search of a kinder, gentler vaccine*. *Science* 296(5573): 1594–1594.
- Foo D.G.W., Macary P.A., Alonso S., Poh C.L. (2008) *Identification of human CD4<sup>+</sup> T-cell epitopes on the VP1 capsid protein of enterovirus 71*. *Viral Immunol.* 21(2): 215–224.
- Fenner F., Hendersn D.A., Arita I., Jezek Z., Ladnyi I.D. (1988) *Smallpox and its eradication*. Geneva: World Health Organization.
- Forouharmehr A., Nassiry M.R. (2015) *B and T-cell epitopes prediction of the P40 antigen for developing mycoplasma agalactiae vaccine using bioinformatic tools*. *Genetics* 3: 3954–3961.
- Giddens W.E., Swango L.J., Henderson J.D., Lewis R.A., Farner D.S., Carlos A., Dolowy W.C. (1971) *Canary pox in sparrows and canaries (Fringillidae) and in weavers (Ploceidae)*. *Vet. Pathol.* 8(3): 260–280.
- Hewitt E.W. (2003) *The MHC class I antigen presentation pathway: strategies for viral immune evasion*. *Immunology* 110(2): 163–169.
- Huang J., Honda W. (2006) *CED: a conformational epitope database*. *BMC Immunol.* 7(1): 7.
- Kennedy R., Poland G.A. (2007) *T-cell epitope discovery for variola and vaccinia viruses*. *Rev. Med. Virol.* 17(2): 93–113.
- Kennedy R.B., Ovsyannikova I.G., Jacobson R.M., Poland G.A. (2009) *The immunology of smallpox vaccines*. *Curr. Opin. Immunol.* 21(3): 314–320.
- Lee H.J., Essani K., Smith G.L. (2001) *The genome sequence of Yaba-like disease virus, a yatapoxvirus*. *Virology* 281(2): 170–192.
- Mack S.J., Tu B., Lazaro A., Yang R., Lancaster A.K., Cao K., Hurley C.K. (2009) *HLA-A, -B, -C, -DRB1 allele and haplotype frequencies distinguish eastern european americans from the general european american population*. *HLA* 73(1): 17–32.
- Moise L., Buller R.M., Schriewer J., Lee J., Frey S.E., Weiner D.B., De Groot, A.S. (2011) *VennVax, a DNA-prime, peptide-boost multi-T-cell epitope poxvirus vaccine, induces protective immunity against vaccinia infection by T cell response alone*. *Vaccine* 29(3): 501–511.
- Molero-Abraham M., Glutting J.P., Flower D.R., Lafuente E.M., Reche P.A. (2015) *EPIPOX: immunoinformatic characterization of the shared T-cell epitome between variola virus and related pathogenic Orthopoxviruses*. *J. Immunol. Res.* 2015; article ID: 738020; doi: 10.1155/2015/738020.
- Osman M.M., ElAmin E.E., Al-Nour M.Y., Alam S.S., Adam R.S., Ahmed A.A., Salih M.A. (2016) *In silico design of epitope based peptide vaccine against virulent strains of HN-Newcastle disease virus (NDV) in poultry species*. *Int. J. Multidiscip. Curr. Res.* 4.
- Pettersen E.F., Goddard T.D., Huang C.C., Couch G.S., Greenblatt D.M., Meng E.C., Ferrin T.E. (2004) *UCSF Chimera – a visualization system for exploratory research and analysis*. *J. Comput. Chem.* 25(13): 1605–1612.
- Patronov A., Doytchinova I. (2013) *T-cell epitope vaccine design by immunoinformatics*. *Open Biol.* 3(1): 120–139.
- Rammensee H.G., Bachmann J., Emmerich N.P.N., Bachor O.A., Stevanović S. (1999) *SYFPEITHI: database for MHC ligands and peptide motifs*. *Immunogenetics* 50(3–4): 213–219.
- Rossmann M.D., Thompson B., Frederick M., Maliarik M., Iannuzzi M.C., Rybicki B.A., Monos D. (2003) *HLA-DRB1\*1101: a significant risk factor for sarcoidosis in blacks and whites*. *Am. J. Hum. Genet.* 73(4): 720–735.
- Soria-Guerra R.E., Nieto-Gomez R., Govea-Alonso D.O., Rosales-Mendoza S. (2015) *An overview of bioinformatics tools for epitope prediction: implications on vaccine development*. *J. Biomed. Inform.* 53(1): 405–414.
- Singh H., Raghava G.P.S. (2001) *ProPred: prediction of HLA-DR binding sites*. *Bioinformatics* 17(12): 1236–1237.
- Terajima M., Cruz J., Raines G., Kilpatrick E.D., Kennedy J.S., Rothman A.L., Ennis F.A. (2003) *Quantitation of CD8<sup>+</sup> T*

- cell responses to newly identified HLA-A\* 0201 – restricted T cell epitopes conserved among vaccinia and variola (Smallpox) viruses.* J. Exp. Med. 197(7): 927–932.
- Tang S.T., Wang M., Lamberth K., Harndahl M., Dziegiel M.H., Claesson M.H., Lund O. (2008) *MHC-I-restricted epitopes conserved among variola and other related orthopoxviruses are recognized by T cells 30 years after vaccination.* Arch. Virol. 153(10): 1833–1844.
- Trott O., Olson A.J. (2010) *AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading.* J. Comput. Chem. 31(2): 455–461.
- Wyatt L.S., Earl P.L., Eller L.A., Moss B. (2004) *Highly attenuated Smallpox vaccine protects mice with and without immune deficiencies against pathogenic vaccinia virus challenge.* Proc. Natl. Acad. Sci. USA 101(13): 4590–4595.
- Wu S., Zhang Y. (2008) *MUSTER: improving protein sequence profile–profile alignments by using multiple sources of structure information.* Proteins Struct. Funct. Bioinf. 72(2): 547–556.
- Zinoviev V.V., Tchikaev N.A., Chertov O.Y., Malygin E.G. (1994) *Identification of the gene encoding vaccinia virus immunodominant protein p35.* Gene 147(2): 209–214.
- Zhang Q., Wang P., Kim Y., Haste-Andersen P., Beaver J., Bourne P.E., Lund O. (2008) *Immune epitope database analysis resource (IEDB-AR).* Nucl. Acids Res. 36(suppl. 2): W513–W518.