International Journal of Molecular and Clinical Microbiology 9(1) (2019) 1082-1089

International Journal of
Molecular and Clinical Microbiology

In silico Design a vaccine Candidate against *Corynebacterium diphtheriae*

Mostafa Norizadeh Tazehkand [1], Orkideh Hajipour [2]

1. Department of Pharmaceutical Biotechnology, Faculty of Pharmacy, Zonguldak Bulent Ecevit University, Zonguldak, Turkey.
2. Department of Biology, Institute of Basic and Applied Sciences, Pamukkale University, Denizli, Turkey

ARTICLE INFO

Article history:
Received 17 April 2019
Accepted 6 May 2019
Available online 1 June 2019

Keywords:
In silico,
Vaccine,
Corynebacterium diphtheriae

ABSTRACT

This study was undertaken to select and analyze epitopes of *Corynebacterium diphtheriae* toxin using online and offline computerized analysis methods, after that affinity of Major Histocompatibility Complex (MHC) epitopes to HLA-DRB1*0101 was analyzed by Hex-protein protein software. The B cell epitopes were predicted via Immune Epitope Database (IEDB) and MHC class II epitopes were predicted by Vaxign software. A physicochemical analysis of candidate vaccine revealed that the designed vaccine has a molecular weight of 59.062 kD. The estimated half-life of candidate vaccine was found to be greater than 30 hours (mammalian reticulocytes, in vitro), greater than 20 hours (yeast, in vivo), and greater than 10 hours (*Escherichia coli*, in vivo). The instability index (II) is computed to be 31.31 (<40), the aliphatic index was found to be 65.80, and the vaccine was considered stable. The grand average of hydropathicity was 0.580, therefore, the vaccine is a hydrophilic protein and probably it interacts with molecules of water. The result of allerTOP, ToxinPred showed that the vaccine is non-allergenic and non-toxic. According to the obtained data from protParam and pepCalc, our designed vaccine is soluble and no transmembrane helix was projected, hence no expression difficulties are anticipated in the development of the vaccine. The membrane helices value of vaccine was 33.64%. The result of protein protein docking analysis showed that maximum affinity of candidate vaccine to HLA-DRB1*0101 with the score of -636.85. The result of this study showed that the candidate vaccine can be stimulate HLA-DRB1*0101.

1. Introduction

Diphtheria is a bacterial infection caused by *Corynebacterium diphtheriae* (William, 2012). *Corynebacterium diphtheriae* after entering the body attaches to the respiratory system and the bacteria synthesizes a toxin that cause weakness, sore throat, fever, and swollen glands in the neck. The bacterial toxin can destroy tissues of the respiratory system. The toxin also gets into the blood and damages the heart, nerves, and kidneys (Centers for Disease Control and Prevention- *Corynebacterium diphtheriae*). Diphtheria is an infection that spreads among humans. The infection can spread through an infected person to any mucous membrane in a new person. The toxic infection most often attacks the lining of the nose and throat (medical news today, 2019).

Diphtheria occurs most often in Sub-Saharan Africa, India, and Indonesia. In 2015, it resulted in 2,100 deaths, down from 8,000 deaths in

*Corresponding author: Dr. Mostafa Norizadeh Tazehkand
Tel: +90 5372891027
E-mail address: mostafa_noorizadeh@yahoo.com

1990. In the United States, 57 cases were reported between 1980 and 2004. Death occurred in 5% to 10% of those affected. In 2016, the countries of worldwide reported about 7,100 cases of diphtheria to the World Health Organization (WHO), but there are likely many more cases (Centers for Disease Control and Prevention- *Corynebacterium diphtheriae*). Only some strains of *Corynebacterium diphtheriae* can synthesize a toxin when infected with a bacteriophage that integrates the toxin-encoding gene into the genome of *Corynebacterium diphtheriae (*Freeman, 1951*).* Diphtheria toxin encodes a protein with 60 Kilo Dalton (protein with 535 amino acid) that composed of two peptide chains (A and B chain), held together by a disulfide bond. Each chain has a role in disease formation. The role of B chain is recognition subunit that gains the toxin entry into the host cell by binding to the EGF-like domain of heparin-binding EGF-like growth factor on the cell surface of host. This signals inters into the cell to internalize the toxin within an endosome by receptor-mediated endocytosis. Inside the endosome, the toxin is spited by a trypsin-like protease. The acidity of the endosome causes B chain to make several pores in the membrane of endosome. The A chain of diphtheria toxin inhibits the synthesis of all proteins in the infected cell by catalyzing of ADP ribosylation of elongation factor EF 2.

EF2 is an important factor to the protein synthesis *(*Freeman et al., 1953*).* Diphtheria vaccine is an effective way for inhibition of diphtheria infection that available in a number of formulations. The diphtheria vaccine is given in 3 or 4 doses along with tetanus and pertussis vaccines which are injected during childhood. In last years, bioinformatics has gathered much scientific consideration. Vaccines established via bioinformatics are safer, more suitable, more effective, and less expensive than other vaccines (Zheng et al., 2017).

Our present study was undertaken to select and analyze epitopes of *Corynebacterium diphtheriae* toxin using different computerized analysis methods. After that affinity of and MHC epitopes to HLA-DRB1*0101 was analyzed by Hex-protein protein software.

2. Materials and Methods

2.1. Tools

This study was administered in silico. Protein sequence of B chain of *Corynebacterium diphtheriae* toxin was obtained from National Center Biotechnology information (NCBI) databank. The 3D structure of HLA-DRB1 was obtained from SWISS-MODEL. Online and offline softwares including of Vaxquery, CDC, VaxiJen, Allertop, Toxinpred, Vaxign, Propred, NCBI, Prabi, 3drefine, Protparam, Prochech, Pepcalc have been used in this work., Iupred2a, Immune Epitope Database (IEDB),25 PEP-FOLD, and Hex Protein Docking were used in this research (Zahroh eat al., 2016; Doytchinova and Flowe, 2007., Vita et al., 2015).

2.2. Methods

Firstly the sequences of *Corynebacterium diphtheriae* toxin (B chain) was obtained from National Center Biotechnology information (NCBI) database (https://www.ncbi.nlm.nih.gov/protein/4AE1_B).

The B cell epitopes were predicted via Immune Epitope Database (IEDB). The software uses an artificial neural network for provide data about potentially B-cell epitopes. In this study epitopes higher than 0.35 threshold were subjected to B-cell epitope prediction.

The epitopes of MHC class II predicted from the selected sequences of *Corynebacterium diphtheriae* toxin (B chain) by Vaxign software. The epitopes were evaluated for their binding affinity with predominant HLA II alleles (P-values < 0.05 were considered significant (Dikhit et al., 2017).

The binding character (with high score) of MHCII and B cell epitopes were taken into attention for the choice of the best epitopes. The antigenicity of selected epitopes was tested with VaxiJen prediction analysis software (Doytchinova and Flower, 2007). The Allergenicity and toxicity of epitopes were analyzed by Allertop and toxinpred softwares. The listed epitopes (B and MHCII epitopes) were linked by Lysine-Lysine linker together and using an immunological adjuvant to build a multi epitopes candidate vaccine. To increase the immunogenicity of candidate vaccine, the amino acid sequence of diphtheria toxin fragment sequence (ID: AAT37555.1) was attached to the N-terminal and C-terminal end of the protein vaccine. Diphtheria toxin fragment is the nontoxic portion of *Corynebacterium*

diphtheriae toxin. A flexible using of a PAPAP linker was used to join adjacent epitopes to candidate vaccine. So, the antigenicity, allergenicity and toxicity of poly-epitope vaccine were tested by Vaxijen, allertop and toxinpred softwares (Dimitrov et al., 2014; Gupta et al., 2013).

The tendency of protein to be soluble in human and *E.coli* on overexpression bacterial host cell was calculated (Shey et al., 2019). Furthermore, half-life, molecular weight, isoelectric pointinstability index, aliphatic index and stability of candidate vaccine was calculated by Paratparam and Papcolc software's (Khan et al., 2018). Parabi server was used to test of potential transmembrane helices in the candidate vaccine structure.

Afterward, 3D structure of candidate vaccine was drawn by SWISS-MODEL. The protein structure was refined using 3D refine software. The server subjected the 5 model to our designed vaccine. The refined models were checked for 3D refine score, GDT-HA score, GTD-TS, RMSD score, and MolProbity, and the best model was chosen. The selected model was investigated by Ramachandran plot analysis with Procheck online software (Shahsavani et al., 2018).

In this research we used the IUPred2A program for detection of disorder region of our protein. The IUPred2A is a program that recognize disordered protein regions ANCHOR2. It is also capable to identify protein regions that do or do not adopt a stable structure depending on the redox state of their environment (Meszaros et al., 2018).

The binding affinity of epitopes to MHC class II (HLA-DRB1*0101) evaluated by HEX protein protein docking software. For this reason molecular structure of HLA-DRB1*0101 in pdb format was obtained from PDB database (http://www.rcsb.org/structure/5V4N). Hex is a communicating molecular graphics software for calculating and exhibiting docking modes of couples of protein and DNA molecules. Hex can calculate protein ligand docking, and it can superpose couples of molecules using only knowledge of their 3D shapes (http://hex.loria .fr/manual800/hex_manual.pdf). For socking step the 3D structure of HLA-DRB1*0101, as a ligand, and the candidate vaccine structure as a receptor were submitted to Hex program. In this research we used human serum albumin (protein structure) as negative control and analyzed the affinity of that protein as receptor to HLA-DRB1*0101 as ligand in Hex software (Amisha et al., 2014).

3. RESULTS

Chain B, Diphtheria Toxin (535aa) sequences of *Corynebacterium diphtheriae* was obtained from NCBI (ACCESSION NO:4AE1_B). This sequence was tested to antigenicity with VaxiJen server. The vaxiJen value of chain B, Diphtheria toxin was 0.605 (above the threshold 0.4). For this reason B chain of Diphtheria Toxin was suitable for our research.

The B cell epitopes were predicted via Immune Epitope Database (IEDB). In this study epitopes higher than 0.35 threshold were subjected to B-cell epitope prediction. These epitopes were tested to antigenicity with VaxiJen. The selected peptide epitopes must also possess antigen as evaluated by VaxiJen (above the threshold 0.4). Some of epitopes from Chain B, Diphtheria Toxin passed these criteria. The predicted B-cell epitopes from candidate toxin which produced the best IEDB and VaxiJen score (Table-1).

The MHC class II epitopes were predicted by vaxign software. The epitopes were evaluated for their binding affinity with predominant HLA II alleles (P-values < 0.05 were considered significant). These epitopes were tested to antigenicity with VaxiJen. The selected peptide epitopes must also possess antigen as evaluated by VaxiJen (above the threshold 0.4). Some of epitopes passed these criteria. The predicted MHCII epitopes which produced the best VaxiJen score, were KSFVMENFS, WAVNVAQVI, and LSLFFEIKS (Table 2-4). These selected epitopes has the strongest affinity to HLA-DRB1. therefore, this HLA type will be used as a model in molecular docking (HEX). The sequence of candidate vaccine is shown in figure 1.

Table 1. The selected of B cell epitopes.

B cell epitopes	Sequence	Length
1	SIGVLGYQKTVDHTKVNSKLSLF	23
2	SYHGTKPGYVDS	12
3	AEGSSSVE	8
4	TRGKRGQDA	9
5	DSETADNLEK	10
6	YNRPAYSPGHKTQPF	15

Table 2. The property of KSFVMENFS epitope.

MHC Class	MHC Allele	p-value
II	HLA-DRB3*01:01	0.000312
II	HLA-DRB5*01:01	0.00208
II	HLA-DRB1*13:02	0.00388
II	HLA-DRB1*03:01	0.00493
II	HLA-DRB1*04:04	0.00517
II	HLA-DRB1*04:03	0.00545
II	HLA-DRB1*07:01	0.00595
II	HLA-DRB1*01:01	0.00778
II	HLA-DRB1*15:01	0.0107
II	HLA-DR4	0.0156
I	HLA-A*30:01	0.019
II	HLA-DPB1*04:01	0.0207
II	HLA-DRB1*09:01	0.0318
II	HLA-DPB1*04:02	0.0383
II	HLA-DRB3*02:02	0.0419

Table 3. The property of WAVNVAQVI epitope.

MHC Class	MHC Allele	p-value
II	HLA-DQA1*01:01/DQB1*05:01	0.00425
II	HLA-DRB1*01:02	0.00655
II	HLA-DPA1*03:01/DPB1*04:02	0.00804
I	HLA-B*53:01	0.0115
I	H-2-Db	0.0149
I	HLA-B*54:01	0.0229
II	HLA-DRB1*15:01	0.0235
I	HLA-B*51:01	0.0287
II	HLA-DRB4*01:01	0.0303
I	HLA-B*39:01	0.0311
I	HLA-B*35:01	0.034
I	HLA-B*15:03	0.0396
I	Patr-B*1301	0.0406
I	HLA-B*27:20	0.0426
I	HLA-A*32:01	0.0463
II	HLA-DRB1*04:02	0.0463
II	HLA-DRB1*03:01	0.0466
II	HLA-DRB1*07:01	0.049

Table 4. The property of LSLFFEIKS epitope.

MHC Class	MHC Allele	p-value
II	HLA-DRB1*04:04	0.0106
II	HLA-DRB1*09:01	0.0109
II	HLA-DRB5*01:01	0.0121
II	HLA-DRB4*01:01	0.0122
II	HLA-DRB1*11:01	0.013
II	HLA-DRB1*04:05	0.0171
II	HLA-DRB1*01:01	0.0178
II	HLA-DRB1*12:01	0.0196
II	HLA-DRB3*01:01	0.0339
II	HLA-DRB1*03:01	0.0423

```
        10         20         30         40         50         60
MARMAMGADD VVDSSKSFVM ENFSSYHGTK PGYVDSIQKG IQKPKSGTQG NYDDDWKGFY

        70         80         90        100        110        120
STDNKYDAAG YSVDNENPLS GKAGGVVKVT YPGLTKVLAL KVDNAETIKK ELGLSLTEPL

       130        140        150        160        170        180
MEQVGTEEFI KRFGDGASRV VLSLPFAEGS SSVEYINNWE QAKALSVELE INFETRGKRG

       190        200        210        220        230        240
QDAMYEYMAQ ACAGNRVRRI MFAPAPSIGV LGYQKTVDHT KVNSKLSLFK KLSLFFEIKS

       250        260        270        280        290        300
KKKSFVMENF SKKALSSLMV AQAIPLVGKK YNRPAYSPGH KTQPFKKTRG KRGQDAKKAE

       310        320        330        340        350        360
GSSSVEKKDS ETADNLEKKK SYHGTKPGYV DSPAPAPMAR MAMGADDVVD SSKSFVMENF

       370        380        390        400        410        420
SSYHGTKPGY VDSIQKGIQK PKSGTQGNYD DDWKGFYSTD NKYDAAGYSV DNENPLSGKA

       430        440        450        460        470        480
GGVVKVTYPG LTKVLALKVD NAETIKKELG LSLTEPLMEQ VGTEEFIKRF GDGASRVVLS

       490        500        510        520        530
LPFAEGSSSV EYINNWEQAK ALSVELEINF ETRGKRGQDA MYEYMAQACA GNRVRRIM
```

Figure 1. Sequence of vaccine candidate.

A physicochemical analysis of candidate vaccine revealed that the designed vaccine has a molecular weight of 59.062 kD. The estimated half-life of candidate vaccine was found to be greater than 30 hours (mammalian reticulocytes, in vitro), greater than 20 hours (yeast, in vivo), and greater than 10 hours (Escherichia coli, in vivo). The instability index (II) is computed to be 31.31 (<40), thus the vaccine was considered stable. The aliphatic index was found to be 65.80, thus the candidate vaccine is probable to be thermostable. The grand average of hydropathicity (GRAVY) was -0.580, thus the vaccine is a hydrophilic protein and probably interact with molecules of water. The allertop analysis showed the non-allergenicity of our vaccine. According to the obtained data from protparam and pepcalc, our designed vaccine is soluble and no transmembrane helix was projected hence no expression difficulties are anticipated in the development of the vaccine. The membrane helices value of vaccine was 33.64% (Figure 2).

The 3d structure of candidate vaccine against *Corynebacterium diphtheriae* toxin was drawn by a SWISS-MODEL online software (Figure 3) and then the model was refined by 3D refine (Figure 4) software. 3D structure of vaccine was suitable due to the lack of good structural templates for homology modeling. The 3d refine score was 7271, GDT-HA score was 0.994, GTD-TS score was 1.000 RMSD score was 0.269, and MolProbity score was 1.163. The selected model was investigated by Ramachandran plot analysis with Procheck online software (Figure 5).

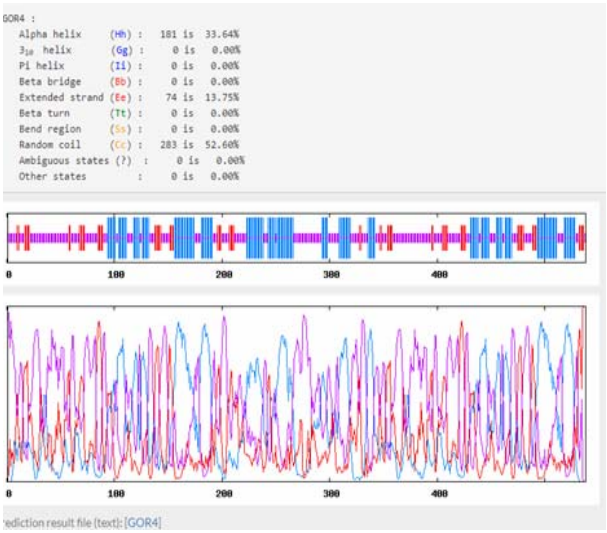

Figure 2. Secondary structure of vaccine candidate.

Figure 3. 3D model of vaccine candidate.

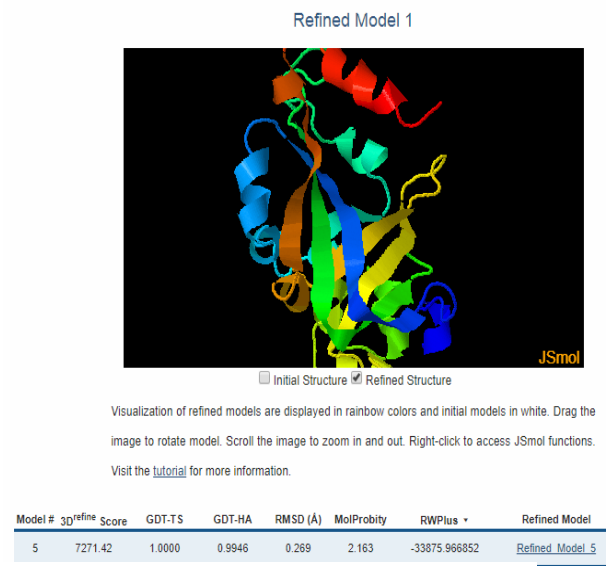

Figure 4. 3D refinement model of vaccine candidate.

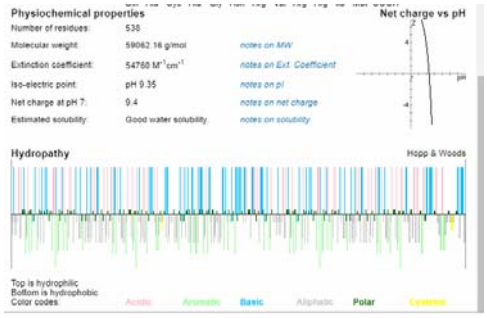

Figure 5. Physiochemical properties of vaccine candidate

Ramachandran plot analysis by procheck revealed that 94.4% of residues are in most favored regions, 5.6% of residues are additional allowed regions and 0% of residues is in disallowed regions (Figure 6). The result of Ramachandran plot analysis supported the high-quality structure of the refined model.

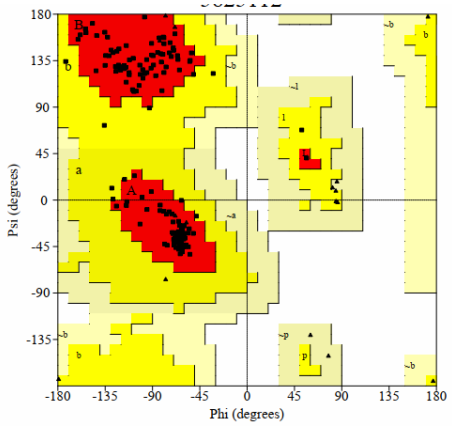

Figure 6. The Ramachandran plot of vaccine candidate.

In this research we used the IUPred2A program for detection of disorder region of our protein. The result obtain from IUPred2A showed that the protein is stable and do not have important disorders (Figure 7).

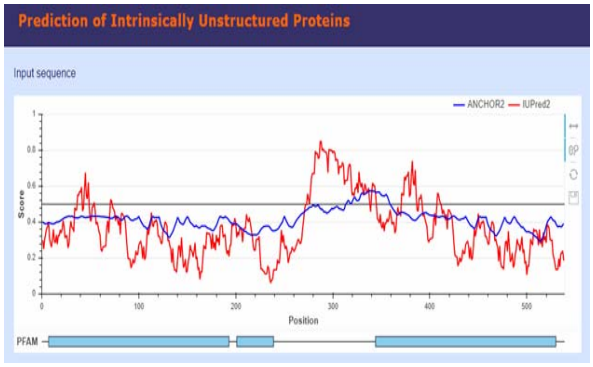

Figure 7. *Prediction of protein disorder using the IUPred web server.*

The binding affinity of epitopes to MHC class II (HLA-DRB1*0101) is tested by HEX protein protein docking software. For molecular docking the 3D structure of HLA-DRB1*0101 used as a ligand and the candidate vaccine structure used as a receptor and then submitted to Hex program. In this work, we used human serum albumin (protein structure) as negative control and analyzed the affinity of that protein as receptor to HLA-DRB1*0101 as ligand in Hex software. The result of protein proten dochong showed that maximum affinity of candidate vaccine to HLA-DRB1*0101 with the score of -636.85 while the affinity of human serum albumin to HLA-DRB1*0101 was -84.16. The result of this study showed that the candidate vaccine can be stimulate HLA-DRB1*0101 (Figure 8).

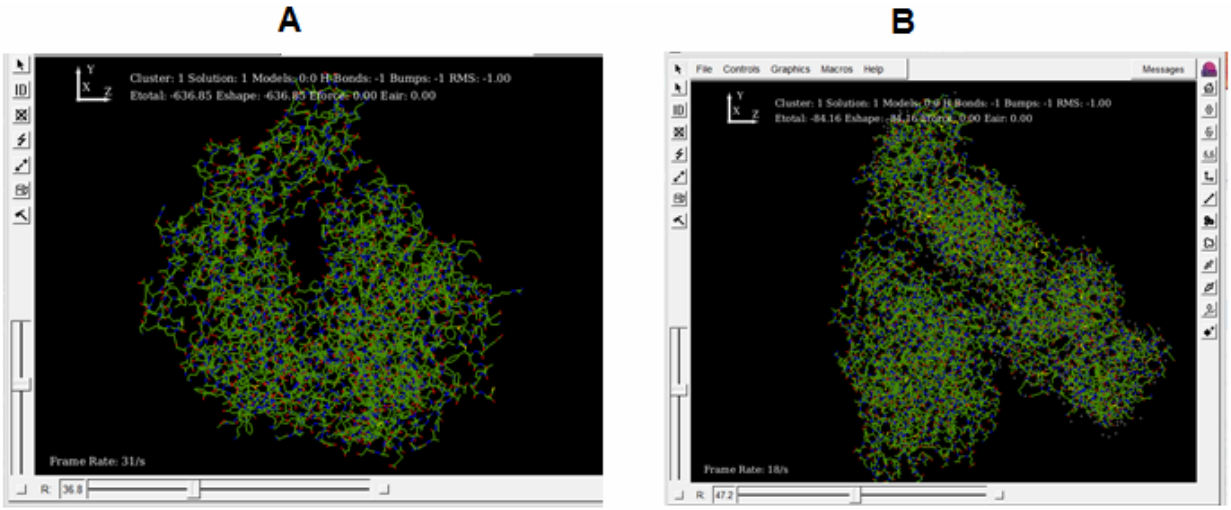

Figure 8. Molecular docking of HLA-DRB1*0101 with candidate vaccine (A) and Human serum albumin (B).

4. Discussion

In this study we used the sequence of B chain toxin of *Corynebacterium diphtheriae* for design and in silico analysis of novel epitope vaccine candidate comprising high score epitopes from antigen. We used immunoinformatical analysis to design and checked the vaccine. HEX protein-protein docking software was used for investigate the affinity of our designed vaccine to a type of MHCII (HLA-DRB1*0101). The solicitation of different software's in vaccine design can importantly stimulate the procedure of vaccine finding and achieve this aim in few time. The peptide sequences having the value above the threshold (above 0.4) can candidate to vaccine. The Peptides having higher scores mean that they are easily recognized by B-cell and T-cells, thus having a higher probability as epitopes. The selected epitopes must also possess antigen as evaluated by VaxiJen (above 0.4). The peptides with antigenic properties are necessary to raise the immune responses.

Cytotoxic T-lymphocyte T-Cell Receptors identify endogenous antigen presented on MHCI, but inflammatory T-Cell Receptors and T- helper identify exogenous antigen presented on MHC II. For this reason in this research we used MHCII epitopes for candidate vaccine against *Corynebacterium diphtheriae* toxin.

These selected epitopes had the strongest affinity to HLA-DRB1*0101. Hence, this HLA type will be used as a model in molecular docking (HEX). A physicochemical analysis of candidate vaccine revealed that the designed vaccine has a molecular weight of 59.062 kD. Proteins that having <110 kD molecular weight could be suitable vaccine candidates (Dar et al., 2019).

The estimated half-life of candidate vaccine was found to be greater than 30 hours (mammalian reticulocytes, in vitro), greater than 20 hours (yeast, in vivo), and greater than 10 hours (*Escherichia coli*, in vivo). The instability index is calculated to be 31.31 that is more less than 40, thus the vaccine was considered as stable. The aliphatic index of our vaccine calculated is high, thus the candidate vaccine is probable to be thermostable. The allergicity analysis of vaccne showed the non-allergenicity of our vaccine, thus the vaccine is not expected to drop harmful allergic reactions in humans. The result of this study Showed that the vaccine candidates have considerable binding with HLA-DRB1*0101.

Conclusion

The candidate vaccine has suitable structural, physiochemical, and immunological properties that may activate humoral and cellular immune responses against of *Corynebacterium diphtheriae* toxin. However, the candidate vaccine could be cloned and expressed at the laboratory and the vaccine experiments using model animals should be performed to confirm the suitability of vaccine against *Corynebacterium diphtheriae*.

Refereces

Dar, H.A., Zaheer, T., Shehroz, M., Ullah, N., Naz, K., Muhammad, S.A., Zhang, T., Amjad, A.A. (2019). Immunoinformatics Aided Design and Evaluation of a Potential Multi-Epitope Vaccine against Klebsiella Pneumoniae. Vaccines (Basel). Sep; 7(3): 88.

Dikhit, M.R., Kumar, A., Das, S., Dehury, B., Rout, A.K., (2017). Identification of Potential MHC Class-II-Restricted Epitopes Derived from Leishmania donovani Antigens by Reverse Vaccinology and Evaluation of Their CD4+ T-Cell Responsiveness against Visceral Leishmaniasis. Front Immunol. 8: 1763.

Dimitrov, I., Bangov, D.R., Flower, Doytchinova, I. (2014). AllerTOP v.2—a server for in silico prediction of allergens. Journal of Molecular Modeling. 20 (6): 2278.

Doytchinova, I.A., Flower, D.R. VaxiJen (2007). a server for prediction of protective antigens, tumour antigens and subunit vaccines. BMC Bioinformatics.8:4.

Freeman Victor, J. (1951). Studies on the Virulence of Bacteriophage-Infected Strains of Corynebacterium Diphtheriae. Journal of Bacteriology. 61 (6): 675–688.

Freeman, V.J., Morse, I.U. (1953). Further Observations on the Change to Virulence of Bacteriophage-Infected Avirulent Strains of Corynebacterium Diphtheriae. Journal of Bacteriology. 63 (3): 407–414.

Gupta, S., Kapoor, P., Chaudhary, K., Gautam, A., Kumar, R., Raghava, G.P. (2013). Open Source Drug Discovery Consortium. In silico approach for predicting toxicity of peptides and proteins. PLoS ONE. 8:e73957.

Jain, A., Tripathi, P., Shrotriya, A., Chaudhary, R., Singh, A. (2015). In silico analysis and modeling of putative T cell epitopes for vaccine design of Toscana virus. 3 Biotech. 5(4): 497–503.

Khan, A., Junaid, M., Kaushik, A.C., Ali, A., Ali, S.S., Mehmood, A., Wei, D.Q. (2018). Computational identification, characterization and validation of potential antigenic peptide vaccines from hrHPVs E6 proteins using immunoinformatics and computational systems biology approaches. PLoS One. 13(5):e0196484.

Mészáros, B., Erdos, G., Dosztányi, Z. (2018). IUPred2A: context-dependent prediction of protein disorder as a function of redox state and protein binding. Nucleic Acids Res. 2;46(W1):W329-W337.

Shahsavani, N., Sheikhha, M.H., Yousefi, H., Sefid, F. (2018). In silico Homology Modeling and Epitope Prediction of NadA as a Potential Vaccine Candidate

in Neisseria meningitides. Int J Mol Cell Med.7(1): 53–68.

Shey, R.A., Ghogomu, S.M., Esoh, K.K., Nebangwa, N.D., Shintouo, C.M. (2019). In-silico design of a multi-epitope vaccine candidate against onchocerciasis and related filarial diseases. Scientific Reports. 9: 4409.

Vita, R., Overton, J.A., Greenbaum, J.A. (2015). The immune epitope database (IEDB) 3.0. Nucleic Acids Res. 43:D405–D412.

William, A. (2012). Diphtheria Epidemiology and Prevention of Vaccine-Preventable Diseases (12 ed.). Public Health Foundation. 9: 215–230.

Zahroh, H., Marup, A., Tambunan, U.S., Parikesit, A.A. (2016). Immunoinformatics Approach in Designing Epitope-based Vaccine Against Meningitis-inducing Bacteria (Streptococcus pneumoniae, Neisseria meningitidis, and Haemophilus influenzae Type b). Drug Target Insights. 1;10:19-29.

Zheng, J., Lin, X., Wang, X., Zheng, L., Lan, S., Jin, S., Ou, Z., Wu, J. (2017). In Silico Analysis of Epitope-Based Vaccine Candidates against Hepatitis B Virus Polymerase Protein. Viruses. 9(5): 112.