ejpmr, 2020,7(6), 94-107

SJIF Impact Factor 6.222

EUROPEAN JOURNAL OF PHARMACEUTICAL AND MEDICAL RESEARCH

www.ejpmr.com

Research Article

ISSN 2394-3211

EJPMR

DESIGNING OF CELL-MEDIATED EPITOPE VACCINE FOR COVID-19 VIRUS SPIKE PROTEIN

Zahra M. Al-Khafaji*[1], Aaisha B. Mahmood[2] and Marium B. Mahmood[3]

[1]Institute of Genetic Engineering and Biotechnology for Postgraduate Studies, University of Baghdad, Iraq.
[2]Ministry of Agriculture, Veterinary Directorate, Baghdad Veterinary Hospital, Al-Dora Hospital, Iraq.
[3]Financial Affairs Dept., Computer Science, University of Baghdad, Iraq.

***Corresponding Author: Zahra M. Al-Khafaji**

Institute of Genetic Engineering and Biotechnology for Postgraduate Studies, University of Baghdad, Iraq.

Article Received on 04/04/2020 **Article Revised on 25/04/2020** **Article Accepted on 15/05/2020**

ABSTRACT

Another attack of coronaviruses comes again, it causes a sever and fatal disease by COVID-19.Immunity is considered as the main way to face viral disease, it becomes more effective if it develops long-lasting activity through cellular immunity. Different short epitopes were derived from COVID-19 spike protein. For cytotoxic cell (CD8+), 7 epitopes were derived 240- **TLLALHRSY**-248, 628- **QLTPTWRVY**-636, 714-**IPTNFTISV**-722, 894-**LQIPFAMQM**-902, 896- **IPFAMQMAY**-904, 898- **FAMQMAYRF**-906, 1065-**VTYVPAQEK**-1073. For helper cells (CD4+) the following epitopes were derived 199-**GYFKIYSKHTPINLV**-213,504 **GYQPYRVVVLSFELL**-518, 505-**YQPYRVVVLSFELLH**-519, 506-**QPYRVVVLSFELLHA**-520, 715-**PTNFTISVTTEILPV**-729, 892-**AALQIPFAMQMAYRF**-906, 893-**ALQIPFAMQMAYRFN**-907, 898-**FAMQMAYRFNGIGVT**-912,895-**QIPFAMQMAYRFNGI**-909. These epitopes were checked for antigenicity, toxicity, allergenicity, similarity to human proteome, they were highly conserved with the reference protein, and with resemble population coverage. Docking studies revealed desired engagement with selected HLA alleles as indicated by binding affinity energy and RMSD values. Theses epitopes are worth to go for real applications.

KEYWORDS: COVID-19, spike protein, coronaviruses, cellular immunity, epitope vaccine Docking MHCs, Iraq.

INTRODUCTION

Another pandemic coronavirus outbreak comes, caused by COVID-19, started in December last year, studies at different levels revealed its similarity to SARS-CoV.[1,2,3] The virus can cause a severe and even fatal disease, it is highly contagious and transmission occurs presumably by airborne droplets and other routes.[4] Human respiratory tract can be infected with a variety of pulmonary viruses due to continuous exposure to the outside environments and high susceptibility of the respiratory mucosa, this could cause that the immune response to be out of control, which may result in tissue damages, functional impairment and reduce lung capacity.[5,6] The recent epidemic studies show that coronaviruses impose a continuous threat to human and economy as they emerge unexpectedly, spread easily leading to catastrophic results.[6]

Viral diseases usually manipulated immunologically, and majority of vaccines are used specially for prophylactic purposes, by interaction of neutralizing antibodies, but the latter alone are often not enough to protect the body against pathogens, and fail to provide long term efficacy and protection against number of viruses, so immune system uses cell-mediated immunity.[7] Two types of T cells act as second line of the adaptive immunity. It is known that T cells immune responses often provide long-lasting immunity[8], since after resolution of the infection, the majority (90-95%) of the effector T cells are eliminated due to programmed cell death and only a small diverse pool of memory cells remains[9,10], therefore combining cellular immunity especially cytotoxic (CTLs/ CD8+) arm with antibodies may provide optimal protective immunity[5], CTLs cells generally play vital role in containment of viral and bacterial infections and lead to pathogens clearance. CTLs cells are MHC I restricted and attack infected cells.[11,12] Moreover, due to its long-lasting memory cells could be created, so its election peptides was explored to be an alternative vaccines besides to other favorite characters.[13] In addition T helper (CD4+), MHC II restricted cells play a regulatory in immune system, as they are mediated the growth and differentiation of both T-effector cells and antibodies producing B lymphocytes.[14,15] Both CTLs and T helper cells practiced their role indirectly and use Major Histocompatibility Complex (MHC) and in human this called human leucocytes antigen (HLAs), there are large number of MHC I and MHC II molecules or alleles.[16]

Both class I and class II restricted T cells carry out their roles in response to T cell epitopes, which are small

linear peptides derived from antigenic protein and displayed on the surface of antigen-presenting cells (APCs) by multiple alleles of MHCs. The diversity of MHC molecules resulted in presence of a wide variety of peptide epitopes to CD4+ and CD8+ T cells, the recognized epitopes derived from boarder range of proteins, in contrast to B cells generally recognize epitopes on the surfaces of proteins.[14]

The ability of epitope-mining tools has fueled the design and development of vaccines using Vaccinomics field which depends on Bioinformatics analytic tools and access to depositories of curated data related to immune reactions, so now it is a common practice to identify the vaccine candidates epitopes using immunoinformatic approach before going to real or wet applications.[14,17,18]

The key ingredient for immunoinformatic-driven vaccines is the initial set of protein sequences that to be likely targets for host immune response. In such cases the surface proteins, secreted proteins, toxins and virulence factors, in addition to proteins highly expressed during growth and replication or stress proteins, all these represented good starts, but this selection should exclude proteins that are highly conserved across species such as house-keeping gene products.[14] The present study aimed to identify epitopes for T cells of previously characterized COVID-19 spike glycoprotein.[19]

MATERIALS AND METHODS

Number of databases and software were used in this study:

NCBI

Used for protein alignment using BLASTp when required.

IEDB database

Used for design T epitopes, (TepiTool) used to predict epitopes with default threshold values.

PDB database

Used to download pdb files of MHC molecules.

VaxiJen server

Used for prediction the antigenicity of epitopes using 0.4 as threshold value for virus group.

AllerCatPro and AllerTop software

Used to estimate the allergenicity of selected appropriate epitopes.

ToxinPred

Used to find the toxicity of the epitopes.

PyRx v.8

Used for docking studies.

PyMOL and Discovery Studio Visualizer used for visualization of docked epitopes.

Chimera software

Used for format manipulation

MarvinSketch software

Used for format manipulation

RESULTS AND DISCUSSION

A cons sequence protein was derived from all the available spike proteins of COVID-19 deposited in NCBI from 11/ Feb to 06/April, 2020 to minimize the genetic differences between spike proteins, this cons sequence was characterized in other study[19], spike glycoprotein associates with greatest number of antigenic epitopes[3], as the virus uses its spike protein as an adhesion factor to facilitate host entry through special receptors.[6] This part of study devoid to predict CD8+ (CTLs) cell epitopes and CD4+ (T helper) cells, and their HLAs molecules.

MHC I

The whole protein was subjected to TepiTool /IEDB using recommended method and 9mer length for all HLA-A and HLA-B available in the database as conservancy over 80%[20], using percentile <1 and binding affinity <200nM, since the peptides with higher affinity are more likely to be selected by MHC molecules and displayed of the cell surface where they can be recognized by T lymphocytes.[14] Another criteria used was choosing the immune proteasome system, since it is believed that this type performs an improve efficiency in antigen presentation.[18, 21] The studies show that there is a complex enzymatic process and the digestion of protein could result in large number of peptides and only 2-2.5% of peptides transported to ER, then to the surface by TAP and the higher values of TAP mean the higher transport rate[18,21,22], and then bind MHC molecules.[14,18]

This resulted in a very large number of CD8+ epitopes. The latter were subjected to estimation of their antigenicity using VaxiJen at 0.4 threshold value as this character means that the epitope can be recognized by T cells, the epitopes were chosen to be away from molecular mimicry to avoid autoimmune reactions by BLASTing each epitope using BLASTp with nr and Ref-Seq databases at Expected value of 0.05, toxicity and allergenicity[23,24,25] were checked using appropriate software, the final selected epitopes are shown in Table 1.

Table 1: CD8+ T cell epitopes and engaged alleles.

start	end	peptide*	IC50	Antigenicity	Proteasome score	TAP	Alleles
240	248	TLLALHRSY	24.2	0.8009	1.27	1.26	HLA-B*15:25
240	248	TLLALHRSY	58.9	0.8009	1.27	1.26	HLA-B*15:02
240	248	TLLALHRSY	78.5	0.8009	1.27	1.26	HLA-B*15:01
240	248	TLLALHRSY	83.4	0.8009	1.27	1.26	HLA-A*29:02
240	248	TLLALHRSY	174.6	0.8009	1.27	1.26	HLA-A*30:02
628	636	QLTPTWRVY	42.9	1.2119	1.72	1.21	HLA-B*15:25
628	636	QLTPTWRVY	79.2	1.2119	1.72	1.21	HLA-B*15:02
628	636	QLTPTWRVY	119.9	1.2119	1.72	1.21	HLA-B*15:01
628	636	QLTPTWRVY	186.2	1.2119	1.72	1.21	HLA-A*30:02
714	722	IPTNFTISV	70.8	0.8820	1.18	0.06	HLA-B*56:01
714	722	IPTNFTISV	135.6	0.8820	1.18	0.06	HLA-B*53:01
714	722	IPTNFTISV	153.6	0.8820	1.18	0.06	HLA-B*35:01
714	722	IPTNFTISV	167.1	0.8820	1.18	0.06	HLA-B*07:02
714	722	IPTNFTISV	179.2	0.8820	1.18	0.06	HLA-B*51:01
894	902	LQIPFAMQM	17.2	1.0680	1.42	1.4	HLA-B*15:25
894	902	LQIPFAMQM	32.9	1.0680	1.42	1.4	HLA-B*15:01
894	902	LQIPFAMQM	39.7	1.0680	1.16	0.26	HLA-A*02:06
894	902	LQIPFAMQM	119.9	1.0680	1.16	0.26	HLA-B*13:01
894	902	LQIPFAMQM	161.9	1.0680	1.16	0.26	HLA-B*15:02
896	904	IPFAMQMAY	2.8	1.4278	1.42	1.17	HLA-B*35:01
896	904	IPFAMQMAY	20.9	1.4278	1.42	1.17	HLA-B*53:01
896	904	IPFAMQMAY	32.3	1.4278	1.42	1.17	HLA-B*15:02
896	904	IPFAMQMAY	82.5	1.4278	1.42	1.17	HLA-A*29:02
896	904	IPFAMQMAY	176.8	1.4278	1.42	1.17	HLA-B*15:25
898	906	FAMQMAYRF	6.3	1.0278	1.45	1.05	HLA-B*35:01
898	906	FAMQMAYRF	11.5	1.0278	1.45	1.05	HLA-B*53:01
898	906	FAMQMAYRF	23.4	1.0278	1.45	1.05	HLA-B*58:01
898	906	FAMQMAYRF	40.9	1.0278	1.45	1.05	HLA-A*23:01
898	906	FAMQMAYRF	43.4	1.0278	1.45	1.05	HLA-B*15:25
898	906	FAMQMAYRF	48.9	1.0278	1.45	1.05	HLA-B*15:02
898	906	FAMQMAYRF	112.9	1.0278	1.45	1.05	HLA-A*29:02
898	906	FAMQMAYRF	123.9	1.0278	1.45	1.05	HLA-B*15:01
898	906	FAMQMAYRF	143	1.0278	1.45	1.05	HLA-A*24:02
1065	1073	VTYVPAQEK	18.5	0.8132	0.94	0.29	HLA-A*11:01
1065	1073	VTYVPAQEK	38	0.8132	0.94	0.29	HLA-A*03:01
1065	1073	VTYVPAQEK	110.7	0.8132	1.5	1.19	HLA-A*30:01
1065	1073	VTYVPAQEK	114.4	0.8132	1.5	1.19	HLA-A*68:01

*Non-toxic, percentile rank <1

The conversancy of the selected epitopes across all the retrieved proteins is shown I Table 2

Table 2: Conservancy of CD8+ T cell epitopes.

Epitope name: start-end	Epitope sequence	Epitope length	Percent of protein sequence matches at identity <= 100%	Minimum identity	Maximum identity
240-248	TLLALHRSY	9	98.86% (87/88)	88.89%	100.00%
628-636	QLTPTWRVY	9	100.00% (88/88)	100.00%	100.00%
714-722	IPTNFTISV	9	100.00% (88/88)	100.00%	100.00%
894-902	LQIPFAMQM	9	100.00% (88/88)	100.00%	100.00%
896-904	IPFAMQMAY	9	100.00% (88/88)	100.00%	100.00%
898-906	FAMQMAYRF	9	100.00% (88/88)	100.00%	100.00%
1065-1073	VTYVPAQEK	9	100.00% (88/88)	100.00%	100.00%

It is known that T cells can only recognize peptide in the context of MHCs (HLAs), so the latter is important component in epitope driven vaccine in the selection of epitopes that binding to MHC[14], the interaction of

epitope and the receptor (TRC) of CD8+ leading to activation, proliferation, differentiation and effector function. On the other hand, HLA genes are the most polymorphic genes in human genome, this besides the restriction phenomenon, resulted in serious problems in vaccine design and population coverage[26,27,28,29], because each allele binds to a particular group of epitopes.[20]

Table 3 shows the frequency of alleles that interact with the selected epitopes.

Table 3: The frequency of alleles for CD8+ T cells.

start	end	peptide	Antigenicity	allele
894	902	LQIPFAMQM	1.0680	HLA-A*02:06
1065	1073	VTYVPAQEK	0.8132	HLA-A*03:01
1065	1073	VTYVPAQEK	0.8132	HLA-A*11:01
898	906	FAMQMAYRF	1.0278	HLA-A*23:01
898	906	FAMQMAYRF	1.0278	HLA-A*24:02
240	248	TLLALHRSY	0.8009	HLA-A*29:02
896	904	IPFAMQMAY	1.4278	HLA-A*29:02
898	906	FAMQMAYRF	1.0278	HLA-A*29:02
1065	1073	VTYVPAQEK	0.8132	HLA-A*30:01
240	248	TLLALHRSY	0.8009	HLA-A*30:02
628	636	QLTPTWRVY	1.2119	HLA-A*30:02
1065	1073	VTYVPAQEK	0.8132	HLA-A*68:01
714	722	IPTNFTISV	0.8820	HLA-B*07:02
894	902	LQIPFAMQM	1.0680	HLA-B*13:01
240	248	TLLALHRSY	0.8009	HLA-B*15:01
628	636	QLTPTWRVY	1.2119	HLA-B*15:01
894	902	LQIPFAMQM	1.0680	HLA-B*15:01
898	906	FAMQMAYRF	1.0278	HLA-B*15:01
240	248	TLLALHRSY	0.8009	HLA-B*15:02
628	636	QLTPTWRVY	1.2119	HLA-B*15:02
894	902	LQIPFAMQM	1.0680	HLA-B*15:02
896	904	IPFAMQMAY	1.4278	HLA-B*15:02
898	906	FAMQMAYRF	1.0278	HLA-B*15:02
240	248	TLLALHRSY	0.8009	HLA-B*15:25
628	636	QLTPTWRVY	1.2119	HLA-B*15:25
894	902	LQIPFAMQM	1.0680	HLA-B*15:25
896	904	IPFAMQMAY	1.4278	HLA-B*15:25
898	906	FAMQMAYRF	1.0278	HLA-B*15:25
714	722	IPTNFTISV	0.8820	HLA-B*35:01
896	904	IPFAMQMAY	1.4278	HLA-B*35:01
898	906	FAMQMAYRF	1.0278	HLA-B*35:01
714	722	IPTNFTISV	0.8820	HLA-B*51:01
714	722	IPTNFTISV	0.8820	HLA-B*53:01
896	904	IPFAMQMAY	1.4278	HLA-B*53:01
898	906	FAMQMAYRF	1.0278	HLA-B*53:01
714	722	IPTNFTISV	0.8820	HLA-B*56:01
898	906	FAMQMAYRF	1.0278	HLA-B*58:01

These epitopes are with promiscuous nature[30,31], the promiscuously binding antigenic epitopes are considered to act against a vast range of immune systems.[32] So the inclusion of promiscuous epitopes i.e. epitopes that are recognized in context of more than one MHC in epitope-driven vaccine may overcome the challenge of genetic restriction of immune system.[14,33,34,35]

Population coverage analysis plays a significant role in the epitope-based vaccine design because of the highly polymorphic nature of MHC molecules[36,37] to ensure the effectiveness of a vaccine in the general population, epitopes must be able to activate the desired immune response for the majority of the target population, and this means that the epitopes would be suitable for use in a large population.[38] Table 4 indicates the population coverage of selected MHC I epitopes.

Table 4: Population coverage for MHC I alleles.

Area	% of Coverage
Oceania	90.66
East Asia	88.57
Northeast Asia	86.99
Southeast Asia	85.89
West Indies	84.86
West Africa	84.29
North America	82.94
South Africa	82.33
Europe	80.66
South Asia	78.87
North Africa	77.43
Central Africa	74.68
East Africa	73.03
Southwest Asia	66.59
South America	62.17
Central America	7.76

East Asia region is considered as one of the hot spots of COVID-19 virus infection. The greater range of geographic areas stands as an additional useful tool for preclinical evaluation of new vaccines.[32]

MHC II

MHC II molecules serve to bind peptides/epitopes encaged with CD4+ T cell receptors to initiate an immune response. Identification of MHC class II restricted peptide epitopes is an important goal in immunological studies, since the activation of CD4+ helper cells is essential for the development of adaptive immunity against pathogens.[39,40,41,42]

A critical step in CD4+ T cell activation is the recognition of epitopes by MHC II molecules.[43] Crystallographic studies revealed that MHC II epitope binding site consists of a groove and several pockets provided by a β-sheet and two α-helices[44,45], the groove is open on both ends so it could accommodate variable number of residues up to 25 residues.[11] The binding groove forms the major pocket which accommodate sidechains of residues. Core region interaction determines the binding affinity and specificity[46], and the immediate flanking residues have been indicated to make contact with the MHC II molecule outside of the binding groove and contribute to MHC-epitope interactions.[47] Computational prediction of MHC II epitopes is theoretical with practical value, since the experimental identification is costly and time consuming. In human, three genes HLA-DR, HLA-DP and HLA-DQ are present, they have large number of alleles which may differ from each other by up to 20 amino acids.

In this study, IEDB recommended method /TepiTool was used for specific top 15 HLA-DR and all the available HLA-DP and HLA-DQ alleles in the database, under percentile <10 and IC50 <200nM. Large number of epitopes were obtained, these were subjected to different steps of filtration and analyses, such as estimation of antigenicity using VaxiJen server at 0.4 threshold value for virus group, estimation similarity with human proteome to avoid autoimmunity problems using BLASTp with nr and Ref-Seq databases at Expected value of 0.05. Toxicity and Allergenicity were checked using appropriate software. Finally 8 epitopes were being satisfied these criteria, shown in Table 5.

Table 5: CD4+ T cell epitopes and engaged allels.

Start	End	Peptide	Antigenicity	IC 50	Allele
199	213	GYFKIYSKHTPINLV	0.9278	79.68	HLA-DRB1*09:01
199	213	GYFKIYSKHTPINLV	0.9278	60.93	HLA-DRB1*11:01
199	213	GYFKIYSKHTPINLV	0.9278	107.67	HLA-DRB1*13:02
199	213	GYFKIYSKHTPINLV	0.9278	92.54	HLA-DRB1*15:01
504	518	GYQPYRVVVLSFELL	1.0740	151.07	HLA-DPA1*01:03/DPB1*02:01
504	518	GYQPYRVVVLSFELL	1.0740	128.38	HLA-DPA1*02:01/DPB1*01:01
504	518	GYQPYRVVVLSFELL	1.0740	151.07	HLA-DPA1*03:01/DPB1*04:02
504	518	GYQPYRVVVLSFELL	1.0740	194.52	HLA-DRB1*04:05
504	518	GYQPYRVVVLSFELL	1.0740	107.53	HLA-DRB1*07:01
504	518	GYQPYRVVVLSFELL	1.0740	189.64	HLA-DRB1*15:01
505	519	YQPYRVVVLSFELLH	0.9711	102.17	HLA-DPA1*01:03/DPB1*02:01
505	519	YQPYRVVVLSFELLH	0.9711	299.25	HLA-DPA1*02:01/DPB1*01:01
505	519	YQPYRVVVLSFELLH	0.9711	127.53	HLA-DPA1*03:01/DPB1*04:02
505	519	YQPYRVVVLSFELLH	0.9711	196.58	HLA-DRB1*04:05
505	519	YQPYRVVVLSFELLH	0.9711	142.47	HLA-DRB1*07:01
506	520	QPYRVVVLSFELLHA	0.9109	90.11	HLA-DPA1*03:01/DPB1*04:02
506	520	QPYRVVVLSFELLHA	0.9109	179.39	HLA-DRB1*07:01
506	520	QPYRVVVLSFELLHA	0.9109	138.75	HLA-DRB1*15:01
506	520	QPYRVVVLSFELLHA	0.9109	187.55	HLA-DRB4*01:01
715	729	PTNFTISVTTEILPV	1.1349	151.86	HLA-DPA1*01/DPB1*04:01
715	729	PTNFTISVTTEILPV	1.1349	151.86	HLA-DRB1*04:01
715	729	PTNFTISVTTEILPV	1.1349	25.32	HLA-DRB1*07:01
715	729	PTNFTISVTTEILPV	1.1349	66.91	HLA-DRB1*09:01

892	906	AALQIPFAMQMAYRF	0.9108	31.62	HLA-DRB1*01:01
892	906	AALQIPFAMQMAYRF	0.9108	171.72	HLA-DRB1*12:01
892	906	AALQIPFAMQMAYRF	0.9108	63.42	HLA-DRB1*15:01
892	906	AALQIPFAMQMAYRF	0.9108	82.35	HLA-DRB4*01:01
892	906	AALQIPFAMQMAYRF	0.9108	31.62	HLA-DRB5*01:01
893	907	ALQIPFAMQMAYRFN	1.0112	13.18	HLA-DRB1*01:01
893	907	ALQIPFAMQMAYRFN	1.0112	159.54	HLA-DRB1*12:01
893	907	ALQIPFAMQMAYRFN	1.0112	72.44	HLA-DRB4*01:01
893	907	ALQIPFAMQMAYRFN	1.0112	23.39	HLA-DRB5*01:01
895	909	QIPFAMQMAYRFNGI	0.9573	12.49	HLA-DRB1*01:01
895	909	QIPFAMQMAYRFNGI	0.9573	180.52	HLA-DRB1*12:01
895	909	QIPFAMQMAYRFNGI	0.9573	51.33	HLA-DRB1*15:01
895	909	QIPFAMQMAYRFNGI	0.9573	70.4	HLA-DRB4*01:01
895	909	QIPFAMQMAYRFNGI	0.9573	21.44	HLA-DRB5*01:01

From the results, it is obvious that the frequent allele belongs to DRB1 and this expected since DRB1 proteins expression level is about five folds greater than those of DRB3, DRB4, DRB5.[48] Conservancy of these epitopes across the whole number of spike protein sequences (88 sequences) from all over the world was 100% as shown in Table 6.

The selected epitopes bind to different alleles range from 4-6 alleles (See Table 5) due to high polymorphism found on the exposed surfaces including the peptide binding groove which resulted in diversity within the population. This means that there are many promiscuous peptides can bind multiple MHC II molecules[49], these promiscuous peptides are a prime target for vaccine and immunotherapy, so many computational tools were developed to facilitate scanning and finding such peptides.[50]

Table 6: Conservancy of CD4+ T cell epitopes.

Epitope #	Epitope name	Epitope sequence	Epitope length	Identity <= 100%	Minimum identity	Maximum identity
1	199-213	GYFKIYSKHTPINLV	15	100.00% (88/88)	100.00 %	100.00%
2	504-518	GYQPYRVVVLSFELL	15	100.00% (88/88)	100.00 %	100.00%
3	505-519	YQPYRVVVLSFELLH	15	100.00% (88/88)	100.00 %	100.00%
4	506-520	QPYRVVVLSFELLHA	15	100.00% (88/88)	100.00 %	100.00%
5	715-729	PTNFTISVTTEILPV	15	100.00% (88/88)	100.00 %	100.00%
6	892-906	AALQIPFAMQMAYRF	15	100.00% (88/88)	100.00 %	100.00%
7	893-907	ALQIPFAMQMAYRFN	15	100.00% (88/88)	100.00 %	100.00%
8	895-909	QIPFAMQMAYRFNGI	15	100.00% (88/88)	100.00 %	100.00%

As with MHC I, population coverage is considered one of the criteria to be looked in selection of epitopes. The IEDB population coverage tool was used for selected epitopes with their interacting alleles, shown in Table 7

Table 7: Population coverage for MHC II alleles.

Area	% of Coverage
North America	77.5
East Asia	76.59
Europe	75.81
South Asia	65.75
West Indies	64.06
North Africa	61.56
East Africa	59.93
Oceania	58.01
West Africa	58.01
Northeast Asia	54.54
Central Africa	54.01
Southeast Asia	51.04
South America	36.83
Southwest Asia	36.4
Central America	25.6
South Africa	7.65

Although the coverage in this case is inferior compared with MHC I (See Table 4), but the promising results that

most combinations of epitopes and alleles cover half the different populations in distinct geographical areas considered by IEDB, this indicates the possibility of using multiple epitopes in one vaccine batch.

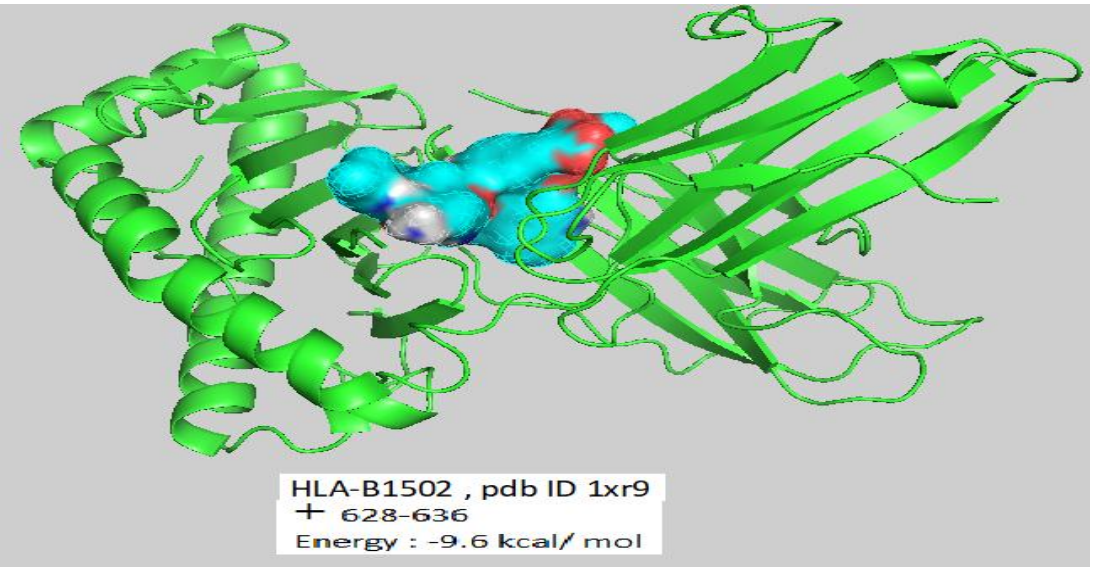
Docking Studies

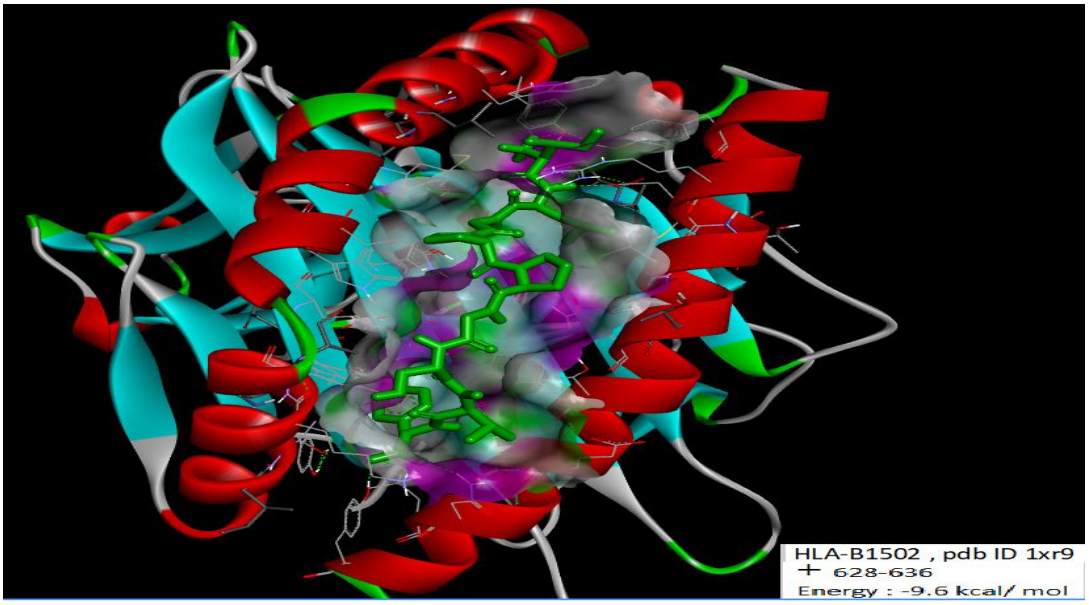
The primary aim of make docking is the prediction of the a binding site of a ligand at a protein receptor surface, then docking and modeling the ligands into the recognize site[20] some epitopes (from MHC I and MHC II) were further tested for binding against HLA molecules using in silico docking technique to verify the binding cleft – epitope interaction.[8] The 3D structure of epitopes were obtained from amino acid sequences using PEP-FOLD online server[51], and the pdb structure of some HLA molecules were obtained from pdb database. The epitope-receptor pairs were docked by PyRx AutoDock Vina[52], the epitope considered as highly flexible ligand to produce a correct docking results[53,34], and in this application the protein (receptor) prepared by removing natural ligands and heteroatoms, addition of polar hydrogen to the structure and ligand torsion are enabled for all rotatable bonds i.e. transforming the format into pdbqt format.[55] The grid box was set to be large enough to accommodate the whole structure of the protein.[38] The best docking results were chosen depending on binding affinity (-Δ G) value, which depicts the binding energy between the protein and the ligands[52] considering that the box of the docking was large enough to let the epitopes to interact with any site of the receptor and considering that PyRx software can dock the epitopes at the same position of crystallography complexes.[13] The results selection was confirmed by RMSD values which were zero for all the selected results i.e. using more stable RMSD in whole docking, since the RMSD value < 3 A^O means high accuracy and those of >3 A^O means low accuracy.[57] On the other hand hydrogen bonding less than 3 A^O is usually considered biologically significant[32], and at the same time the docking or the interaction to bind or dock in the groove carry out via multiple contacts, with continuous hydrogen bonds and salt bridge anchors supporting their potential in generating immune responses.[58]

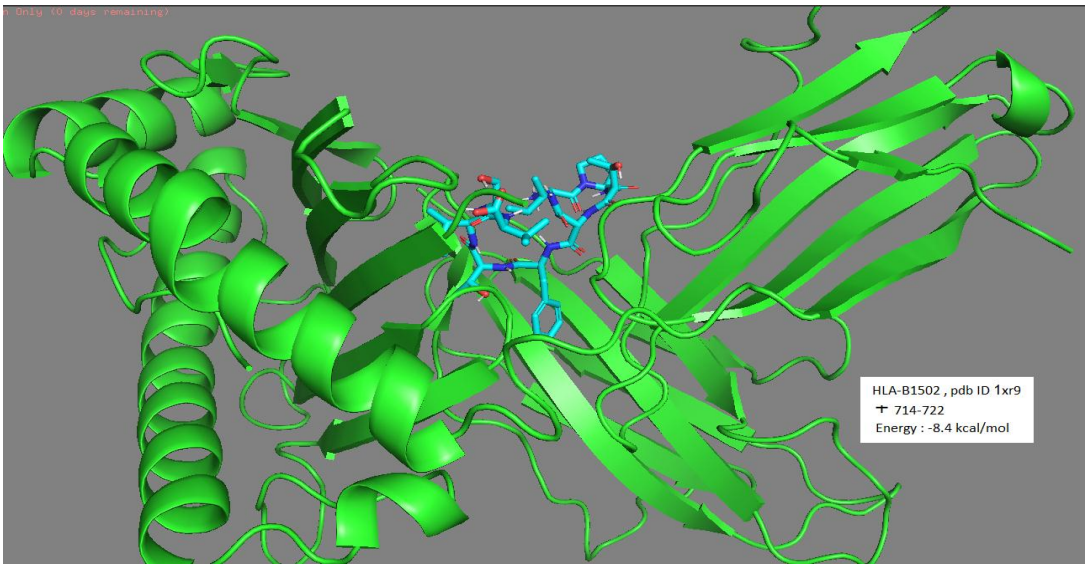
Anyway, in this study, docking was performed for promiscuous peptides which are engaged with more frequent allele, and sometimes using the more frequent alleles.

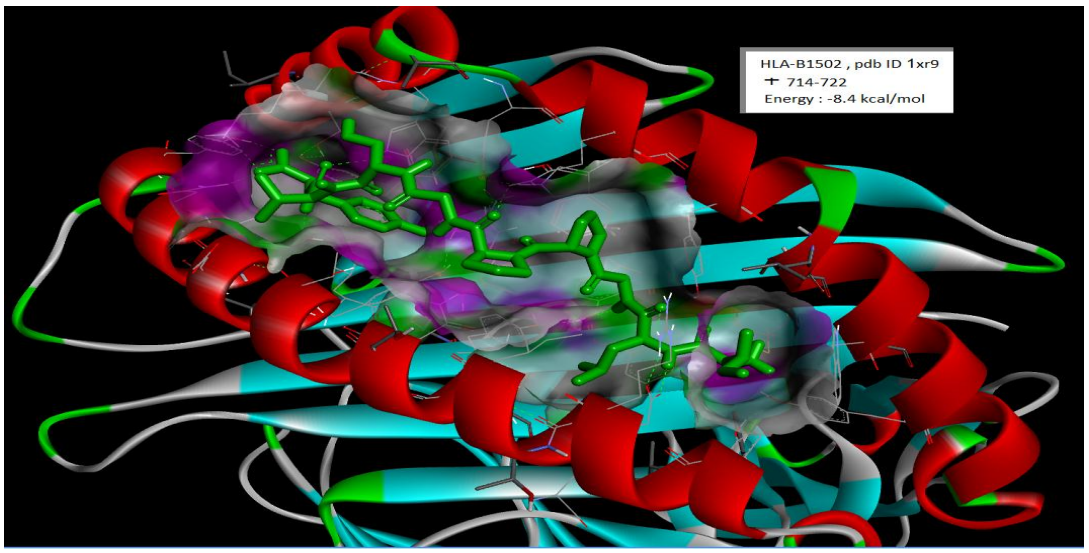
MHC I

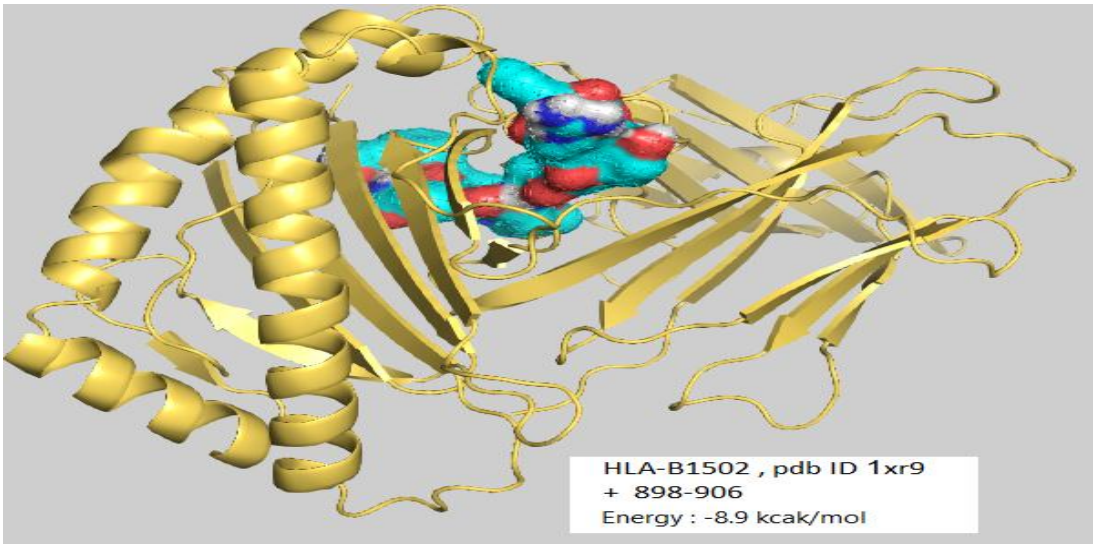
Allele	Epitope	Binding affinity kcal/mol
HLA-B1502 , pdb ID 1xr9	240-248	-8.6
HLA-B1502 , pdb ID 1xr9	628-636	-9.6
HLA-B1502 , pdb ID 1xr9	714-722	-8.4
HLA-B1502 , pdb ID 1xr9	894-902	-8.0
HLA-B1502 , pdb ID 1xr9	896-904	-8.2
HLA-B1502 , pdb ID 1xr9	898-906	-8.9
HLA-B1502 , pdb ID 1xr9	1065-1073	-8.9

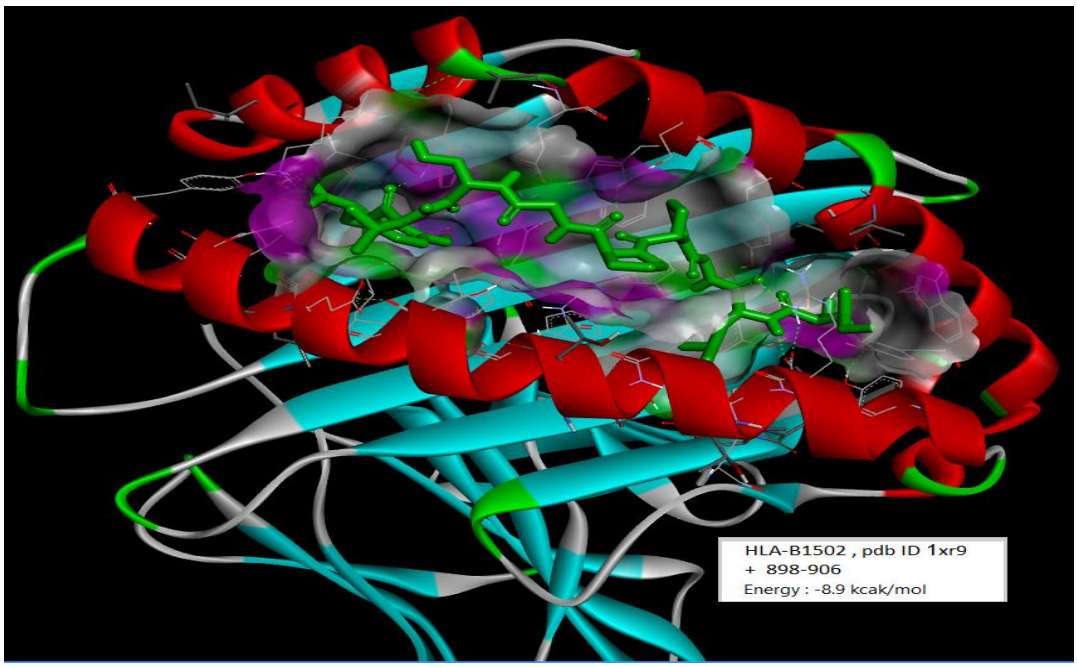



HLA-B1502 , pdb ID 1xr9
+ 628-636
Energy : -9.6 kcal/ mol


HLA-B1502 , pdb ID 1xr9
+ 714-722
Energy : -8.4 kcal/mol

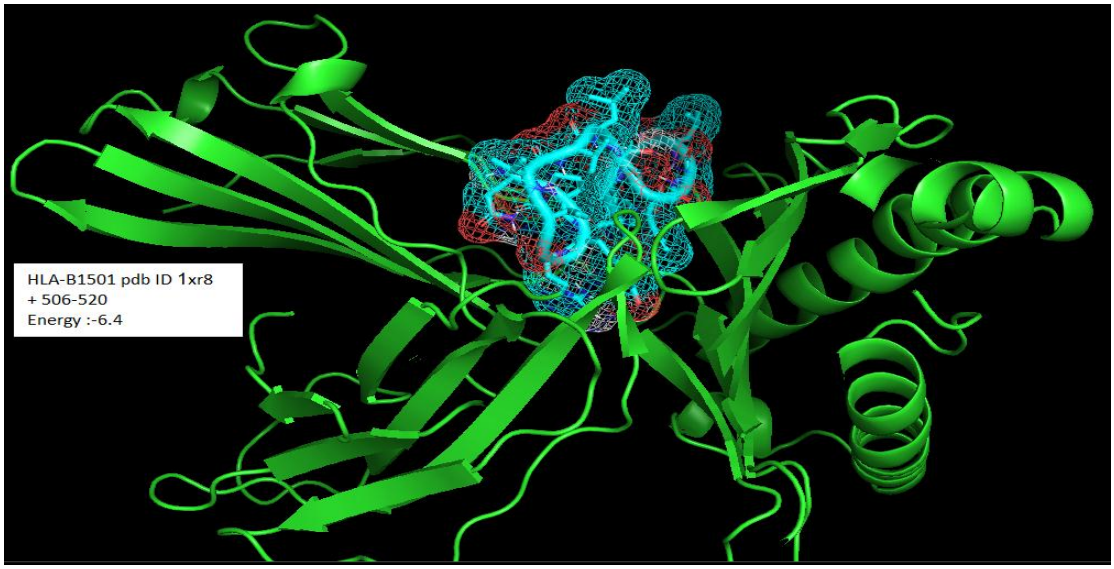

HLA-B1502 , pdb ID 1xr9
+ 714-722
Energy : -8.4 kcal/mol

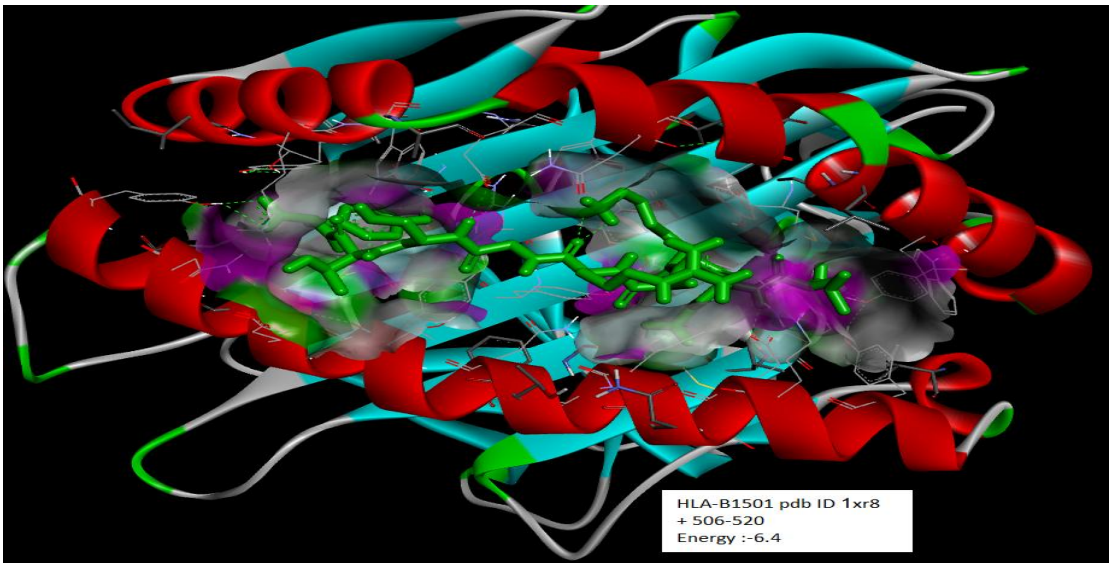


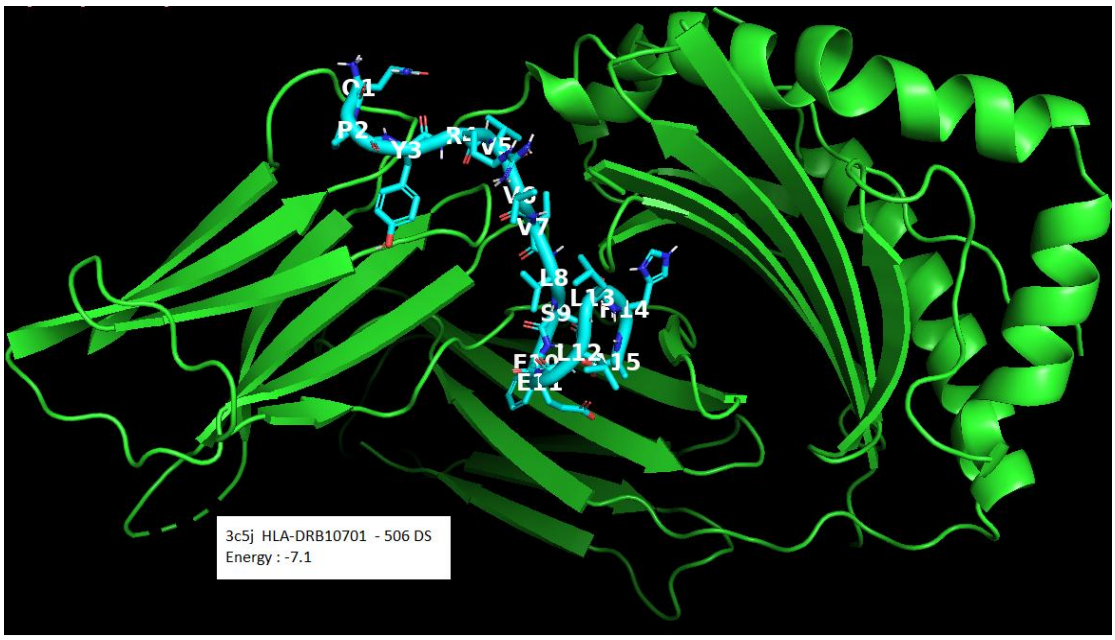


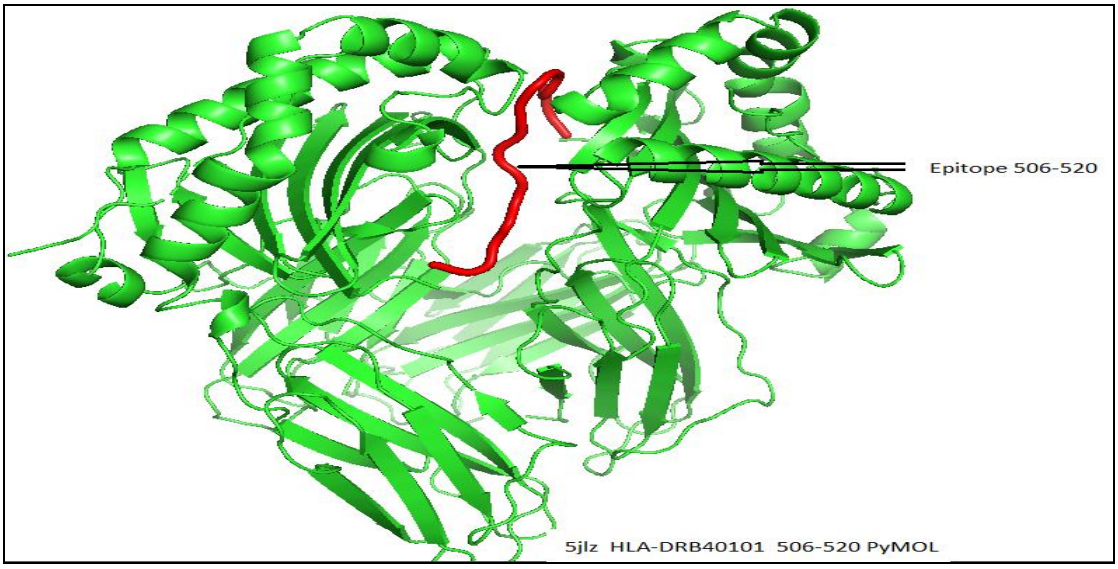
MHC II

Allele	Epitope	Binding affinity kcal/mol
HLA-B1501 pdb ID 1xr8	506-520	-6.4
HLA-DRB10701 pdb ID 3c5j	715-729	-6.6
DRB1 0101 pdb ID 1aqd	892906	-7.1
HLA-DRB10701 pdb ID 3c5j	506-520	-7.1
HLA-DRB40101 pdb ID 5jlz	892-906	-7.5
DRB1 0101 pdb ID 1aqd	506-520	-7.6
HLA-DRB10405 pdb ID 4is6	504-518	-7.7
HLA-B1501 pdb ID 1xr8	199-213	-8.0
HLA-DRB40101 pdb ID 5jlz	506-520	-9.1

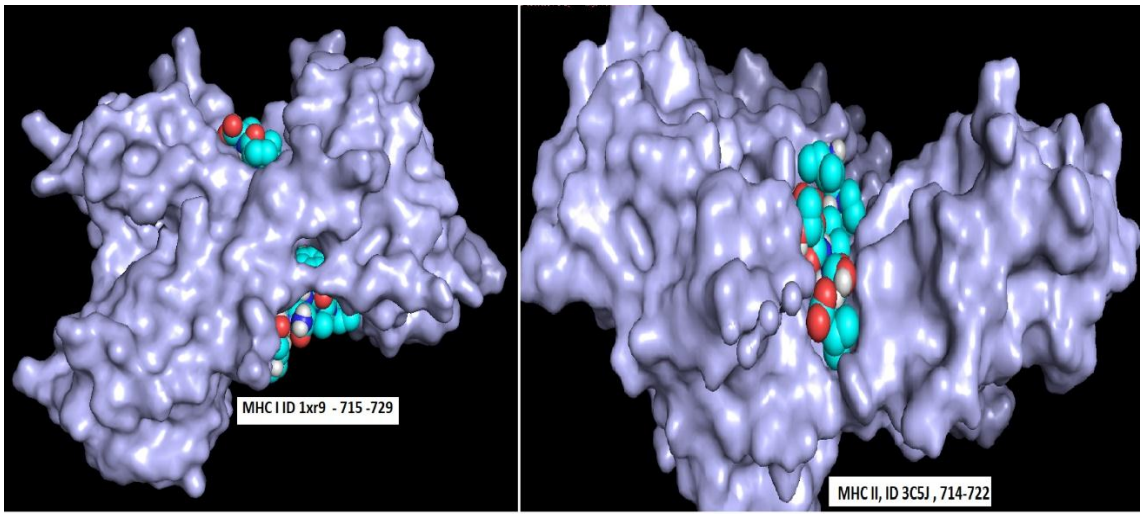

HLA-B1501 pdb ID 1xr8
+ 506-520
Energy :-6.4


HLA-B1501 pdb ID 1xr8
+ 506-520
Energy :-6.4


3c5j HLA-DRB10701 - 506 DS
Energy : -7.1



It has been noted that there is across epitopes i.e. epitope could engaged with MHC I allele and MHC II alleles as for the epitopes 714- IPTNFTISV-722 and 715-PTNFTISVTTEILPV-729, such epitope can engaged with MHC I of CTLs and MHC II of CD4+ T helper cells and represent strategic choice.



The designed ideal vaccine for respiratory viruses should include the cellular and humoral neutralizing antibodies, that because classical antibody- based vaccine are often poor inducer of T cell responses, so including small protein fragments (epitopes) in vaccine which can be presented by MHC molecules to CD4+ and CD8+ T cells, this will lead to specific T cell responses[59] in that CD8+ helping in clearing out the infection[5], while CD4+ T cell functions as helper cells can direct the activity of other immune cells against a viral threats by releasing specific mediators[16] and are critically important to the development of memory B cells and memory CTLs responses.[7,11]

To increase the population coverage, this can be done by using a peptide cocktail composed from different immunogenic peptides, so the immune epitope- based vaccine must contain enough epitopes restricted by supertype HLA to induce broad responses human population.[14] In addition, inclusion of diverse or nonidentical epitopes will improve paracrine effect (cooperation) away from competition found in similar epitopes.[60]

Finally other approaches can be practiced against COVID-19 such as finding inhibitors for most important virulence factor (i.e. spike protein), or using antimicrobial peptides[61], or using available knowledge to targeting and controlling the cytokines production and inflammatory responses.[6]

REFERENCES

1. Sigrist CA, Bridge A, Le Mercier P. A potential role for integrins in host cell entry by SARS-CoV-2. Antiviral Res, 2020; 177, 104759: 1-3.
2. Jalavaa K. First respiratory transmitted food borne outbreak? Int J Hyg Environ Health, 2020; 226, 113490: 1-3.
3. Grifoni A, Sidney J, Zhang Y, Scheuermann R, Peters B, Sette A1 et al., A Sequence homology and Bioinformatic approach can predict candidate targets for Immune Responses to SARS-CoV-2, Cell Host Microbe, 2020; 27: 671–80.
4. Zhang W, Du R, Li B, Zheng XS, Yang XL, Hu B. Molecular and serological investigation of 2019 nCoV infected patients: implication of multiple shedding routes. Emerg Microb Infect, 2020; 9: 386–9.
5. Schmidt ME, Varga SM. The CD8 T cell response to respiratory virus infections:1-12. Front. Immunol, 2018; 9: 678.
6. Geng Li, Fan Y, Lai Y, Han T, Li Z Zhou P, et al. Coronavirus infections and immune responses. J Med Virol, 2020; 1–9.
7. Panagioti E, Klenerman P, Lee LN, van der Burg SH, Arens R. Features of effective T Cell-inducing vaccines against chronic viral infections. Front Immunol, 2018; 9, 276: 1-11.
8. Oany AR, Emran A, Jyoti TP. Design of an epitope-based peptide vaccine against spike protein of human coronavirus: an in silico approach. Drug Des. Dev. Ther, 2014; 8: 1139–49.
9. Zhang M, Byrne S, Liu N, Wang Y, Oxenius A, Ashton-Rickardt PG. Differential survival of cytotoxic T cells and memory cell precursors. J Immunol, 2007; 178: 3483–91.
10. Gerlach C, van Heijst JW, Schumacher TN. The descent of memory T cells. Ann N Y Acad Sci, 2011; 1217: 139–53.
11. He Y, Rappuoli R, DeGroot A, Chen RT. Emerging vaccine informatics. J Biomed Biotechnol. 2010, Article ID 218590, 26 pages.
12. Plotnicky H Cyblat-Chanal D, Aubry JP, Derouet F, Klinguer-Hamour C, Beck A. et al. The immunodominant influenza matrix T cell epitope recognized in human induces influenza protection in HLA-A2/K transgenic mice. Virology, 2003; 309: 320–9.
13. Mohammadi E, Pirkhezranian Z, Monhemi H, Razmyar J, Tahmoorespur M, Sekhavti MH. Epitope characterization, docking and molecular dynamic simulation studies on two main immunogenic Canarypox virus proteins. Biotecnol Apl, 2019; 36: 1211-8.
14. De Groot AS, Ardito M, Tassone R, Knopf P, Moise L, Martin W. Tools for vaccine design: prediction and validation of highly immunogenic and conserved class II epitopes and development of epitope-driven vaccines. In" Development of Vaccines: From Discovery to Clinical Testing ", First Edition. Edited by Manmohan Singh and Indresh K. Srivastava. © 2011 John Wiley & Sons.
15. Ahlers JD, Belyakov IM, Thomas EK, Berzofsky JA., High-affinity T helper epitope induces complementary helper and APC polarization, increased CTL, and protection against viral infection. J Clin Invest, 2001; 108: 1677–85.
16. Bojin F, Gavriliuc O, Margineanu M, Paunescu V. Design of an epitope-based synthetic long peptide vaccine to counteract the novel china coronavirus (2019-nCoV). *Preprints* 2020, 2020020102.
17. Gupta SK, Singh A, Srivastava M, Gupta SK, Akhoon. In silico DNA vaccine designing against human papillomavirus (HPV) causing cervical cancer. Vaccine, 2009; 28: 120–31.
18. Martini S Nielsen M, Peters B1, Sette A. The Immune Epitope Database and Analysis Resource Program 2003–2018: reflections and outlook. Immunogenetics, 2020; 72: 57–76.
19. Al-Khafaji ZM, Mahmood AB, Mahmood MB. Evolution of COVID-19 virus and designing B epitope vaccine. 2020; Eur J Pharm Med Res. 2020; 7/6. In press.
20. Panahi HA, Bolhassani A, Javadi G, Noormohammadi Z. A comprehensive in silico analysis for identification of therapeutic epitopes in HPV16, 18, 31 and 45 oncoproteins. PLoS ONE, 2018; 13: e0205933.
21. Tenzer S, Peters B, Bulik S, Schoor O, Lemmel C, Schatz M, et al. Modeling the MHC class I pathway

by combining predictions of proteasomal cleavage, TAP transport and MHC class I binding. Cell Mol Life Sci, 2005; 62: 1025–37.
22. Peters B, Bulik S, Tampe R, Van Endert PM, Holzhutter HG. Identifying MHC class I epitopes by predicting the TAP transport efficiency of epitope precursors. J Immunol, 2003; 171: 1741–9.
23. Jacob CO, Leitner M, Zamir A, Salomon D, Arnon R. Priming immunization against cholera toxin and *E. coli* heat-labile toxin by a cholera toxin short peptide-beta galactosidase hybrid synthesized in *E. coli*. EMBO J, 1985; 4: 3339-43.
24. Maurer-Stroh S, Krutz N, Kern NS, Gunalan V, Nguyen MN, Limviphuvadh V. et al. AllerCatPro—prediction of protein allergenicity potential from the protein sequence. Bioinformatics, 2019; 35: 3020–7.
25. Dimitrov I, Flower DR, Doytchinova I. AllerTOP - a server for *in silico* prediction of allergens. BMC Bioinform, 2013; 14(Suppl 6): 1-9.
26. Paris R, Bejrachandra S, Thongcharoen P, Nitayaphan S, Pitisuttithum P, Sambor A, et al. HLA class II restriction of HIV-1 clade-specific neutralizing antibody responses in ethnic Thai recipients of the RV144 prime-boost vaccine combination of ALVAC-HIV and AIDSVAX B/E. Vaccine, 2012; 30: 832–6.
27. Singh SP, Mishra BN. Major Histocompatibility Complex linked databases and prediction tools for designing vaccines. Hum. Immunol, 2016; 77: 295–306.
28. Abbas AK, Lichtman AH, Pillai S. Cellular and Molecular Immunology. 8TH edition: Elsevier Health Sciences, 2014.
29. Bui HH, Sidney J, Dinh K, Southwood S, Newman MJ, Sette A. Predicting population coverage of T cell epitope-based diagnostics and vaccines. BMC Bioinform, 2006; 7,153: 1-5.
30. Gasteiger E C, Hoogl CA, Gattikeretal A. Protein Identification and analysis tools on the ExPASy Server, Springer, Berlin, Germany, 2005.
31. Khan S, Ranganathan S. pDOCK: a new technique for rapid and accurate docking of peptide ligands to Major Histocompatibility Complexes," Immunome Res, 2010; 6(Suppl 1): S2: 1-16.
32. Hossain S, Azad A, Chowdhury PA, Wakayama M. Computational identification and characterization of a promiscuous T-Cell epitope on the extracellular protein 85B of *Mycobacterium* spp. for peptide-based subunit vaccine design. Biomed Res Int. 2017; Article ID 4826030, 14 pages.
33. McMurry JA, Sbai H, Gennaro ML, Carter EJ, Martin W, De Groot AS. Analyzing *Mycobacterium tuberculosis* proteomes for candidate vaccine epitopes. Tuberculosis (Edinb), 2005; 85: 95–105.
34. Paina-Bordignon P, Tan A, Termijtelen A, Demotz S, Corradin G, Lanzavecchia A. Universally immunogenic T cell epitopes: Promiscuous recognition by T cells. Eur. J. Immunol, 1989; 19: 2237-42.
35. De Groot AS, Hosmalin CM, Hughes A, Hughes D, Barnd CW, Hendrix R, et al. Human immunodeficiency virus reverse transcriptase T helper epitopes identified in mice and humans: correlation with a cytotoxic T cell epitope. J. Inf. Dis, 1991; 164: 1058–65.
36. Reche PA, Reinherz EL. Sequence variability analysis of human class I and class II MHC molecules: functional and structural correlates of amino acid polymorphisms. J Mol Biol, 2003; 331: 623–41.
37. Stern L J, Wiley DC. Antigenic peptide binding by class I and class II histocompatibility proteins. Structure, 1994; 2: 245–51.
38. Prasasty VD, Grazzolie K, Rosmalena R, Yazid F, Ivan FX Sinaga E. Peptide-based subunit vaccine design of T- and B-cells multi-epitopes against Zika virus using immunoinformatics approaches. Microorganisms, 2019; 7, 226: 1-27.
39. Dranoff G, Jaffee E, Lazenby A, Golumbek P, Levitsky H, et al. Vaccination with irradiated tumor cells engineered to secrete murine granulocyte-macrophage colony-stimulating factor stimulates potent, specific, and long-lasting anti-tumor immunity. Proc Natl Acad Sci U S A, 1993; 90: 3539–43.
40. Jenkins MK, Khoruts A, Ingulli E, Mueller DL, McSorley SJ, et al. In vivo activation of antigen-specific CD4 T cells. Annu Rev Immunol, 2001; 19: 23–45.
41. Germain RN. MHC-dependent antigen processing and peptide presentation: providing ligands for T lymphocyte activation. Cell, 1994; 76: 287–99.
42. Rosenberg ES, Billingsley JM, Caliendo AM, Boswell SL, Sax PE, et al. Vigorous HIV-1-specific CD4+ T cell responses associated with control of viremia. Science, 1997; 278: 1447–50.
43. Rudolph MG, Stanfield RL, Wilson IA. How TCRs bind MHCs, peptides, and coreceptors. Annu Rev Immunol, 2006; 24: 419–66.
44. Stern L J, Brown JH, Jardetzky TS, Gorga JC, Urban RG, et al. Crystal structure of the human class II MHC protein HLA-DR1 complexed with an influenza virus peptide. Nature, 1994; 368: 215–21.
45. Zhu Y, Rudensky AY, Corper AL, Teyton L, Wilson IA. Crystal structure of MHC class II I-Ab in complex with a human CLIP peptide: prediction of an I-Ab peptide-binding motif. J Mol Biol, 2003; 326: 1157–74.
46. Jones EY, Fugger L, Strominger JL, Siebold C. MHC class II proteins and disease: a structural perspective. Nat Rev Immunol, 2006; 6: 271–82.
47. Godkin AJ, Smith KJ, Willis A, Tejada-Simon MV, Zhang J, et al. Naturally processed HLA class II peptides reveal highly conserved immunogenic flanking region sequence preferences that reflect antigen processing rather than peptide-MHC interactions. J Immunol, 2001; 166: 6720–7.
48. Contini S, Pallante M, Vejbaesya S, Park M, Chierakul N, Kim H. et al. S. A model of phenotypic

susceptibility to tuberculosis: Deficient in silico selection of *Mycobacterium tuberculosis* epitopes by HLA alleles. Sarcoidosis Vasc Dif, 2008; 25: 21-8.
49. Consogno G, Manici S, Facchinetti V, Bachi A, Hammer J, et al. Identification of immunodominant regions among promiscuous HLA-DR restricted CD4+ T-cell epitopes on the tumor antigen MAGE-3. Blood, 2003; 101: 1038–44.
50. Zhang GL, Khan AM, Srinivasan KN, August JT, Brusic V. MULTIPRED: a computational system for prediction of promiscuous HLA binding peptides. Nucleic Acids Res, 2005; 33: W172-9.
51. Thevenet P, Shen Y, Maupetit J, Guyon F, Derreumaux P, Tuffery P. PEP-FOLD: an updated de novo structure prediction server for both linear and disulfide bonded cyclic peptides. Nucleic Acids Res, 2012; 40: W288-93.
52. Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem, 2010; 31: 455-61.
53. Hauser AS, Windshügel B. LEADS-PEP: a benchmark data set for assessment of peptide docking performance. J Chem Inf Model, 2016; 56: 188–200.
54. Rentzsch R, Renard BY. Docking small peptides remains a great challenge: An assessment using AutoDock Vina. Brief. Bioinform, 2015; 16: 1045–56.
55. Kumar, S. Drug and Vaccine Design against Novel Coronavirus (2019-nCoV) Spike Protein through Computational Approach. *Preprints* 2020, 2020020071 (doi: 10.20944/preprints202002.0071.v1).
56. Shityakov S, Förster C. In silico predictive model to determine vector-mediated transport properties for the blood–brain barrier choline transporter. Adv. Appl. Bioinform. Chem, 2014; 7: 23–36.
57. Blaszczyk M, Kurcinski M, Kouza M, Wieteska L, Debinski A, Kolinski A, et al. Modeling of protein–peptide interactions using the CABS-dock web server for binding site search and flexible docking. Methods, 2016; 93: 72–83.
58. Baruah V, Bose S. Immunoinformatics-aided identification of T cell and B cell epitopes in the surface glycoprotein of 2019-nCoV. J Med Virol. J Med Virol, 2020; 92: 495-500.
59. Rosendahl Huber S, van Beek J, de Jonge J, Luytjes W, van Baarle, D. T Cell responses to viral infections – opportunities for peptide vaccination. Front Immunol, 2014; article 171, 5: 1-12.
60. Creusot RJ, Thomsen LL, Tite JP, Chain BM. Local cooperation dominates over competition between CD4+T-cells of different antigen/MHC specificity. J Immunol, 2003; 171: 240–6.
61. Prompetchara E, Ketloy C, Palaga T. Immune responses in COVID-19 and potential vaccines: Lessons learned from SARS and MERS epidemic. Asian Pac J Allergy Immunol, 2020; 38: 1-9.