Advanced Research in Microbial Metabolites and Technology
ARMMT

An immunoinformatic approach for a therapeutic epitope-based peptide vaccine design by targeting HPV16-E6: *in silico* study

Bahareh Bahmani[1], Zahra Amini-bayat[1,2]*, Mohammad Mehdi Ranjbar[3], Mohammad Reza Masjedi[2], Amir-Hassan Zarnani[4,5]

[1] *Department of Biotechnology, Iranian Research Organization for Science and Technology (IROST), Tehran, Iran*
[2] *Cancer Control Research Center, Cancer Control Foundation, Iran University of Medical Sciences, Tehran, Iran*
[3] *Department of Virology, Razi Vaccine and Serum Research Institute, Agricultural Research, Education and Extension Organization (AREEO), Tehran, Iran*
[4] *Department of Immunology, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran*
[5] *Reproductive Immunology Research Center, Avicenna Research Institute, ACECR, Tehran, Iran*

Article Info

Received 12/09/2022
Received in revised form 30/11/2022
Accepted 27/12/2022

Keywords:
Bioinformatics approaches,
Cervical cancer,
Epitope design,
HPV-E6,
Therapeutic vaccine

Abstract

Cervical cancer is one of the leading causes of death worldwide, causing approximately 500,000 new cases and 250,000 cancer deaths each year. Persistent infection with high-risk human papillomaviruses (HPVs), particularly type 16, is the primary cause of cervical cancer development and maintenance among women worldwide. The E6 therapeutic vaccines can induce strong anti-tumor T cell-mediated immune responses, such as cytotoxic T lymphocytes, that play vital roles in current therapeutic vaccine development. In our study, bioinformatics approaches and in silico analyzes, such as protein sequence retrieval, identification of conserved regions, drawing of pedigrees, prediction of T-cell epitopes, calculation of population coverage of predicted epitopes, and molecular docking, were used to predict the major histocompatibility complex (MHC) Class I and Class II T cell epitopes of HPV16 E6. Taking into account the scores from different steps, six CD8+ T cells and three CD4+ epitopes were selected. The fusion of the selected epitopes created a universal potential vaccine with a population coverage of 86.41%. The population coverage was obtained by evaluating the potential of these epitopes to elicit innate and acquired immunity. These theoretically confirmed peptides could be employed in a poly-epitope construct as a candidate vaccine for further analyses. Also, these results provide new insights into therapeutic vaccine development.

1. Introduction

Worldwide, cervical cancer has become one of the leading causes of death among women, predominantly in developing countries. The prevalence of cervical cancer varies significantly across different regions based on the countries' development. High incidence rates are reported in South-Central and South East Asia, Africa, and Latin America. Also, low rates are observed in Europe and North America. Epidemiologic studies over the years have emphasized the key role of persistent infection of high-risk human papillomaviruses (HPVs) genotypes in the growth and maintenance of cervical cancer among

*Corresponding author. Tel: *+982157416815*
E-mail address: Amini-bayat@irost.org, Zahra_aminibaiat@yahoo.com.
DOI: 10.22104/ARMMT.2022.1232

women worldwide (Bosch, Lorincz, Muñoz, Meijer, & Shah, 2002).

HPVs are the small icosahedral DNA viruses that belong to the papillomavirus family. Over 170 types of HPVs are classified as low-and high risk according to tumorigenic potential. The high-risk or cancer-causing HPV types include 16, 18, 31, 33, 35, 39, 45, 52, 58, 66, and 68 (Ahmed, Bensumaidea, & Ashankyty, 2015). Among high-risk HPVs, HPV 16 is the most carcinogenic genotype, implicated in 60% of cervical cancer cases; it has also been detected in vaginal, vulvar, penile cancers, anal, and no-genital oropharyngeal head and neck cancers. Similar to E7, E6 can induce and maintain a malignant state in infected cells due to its role in tumorigenesis through the downregulation of the p53 tumor suppressor protein (Boda et al., 2018).

Although HPV prophylactic vaccines have been available since 2006 and induce protective antibodies that prevent the HPV virus from entering into the cervical cells (Stanley, 2007), unfortunately, these vaccines show no benefit in women with persistent infection and do not exert a therapeutic effect against HPV mediated transformed lesions and neoplasia (Lin, Doolan, Hung, & Wu, 2010). Therefore, humans require therapeutic HPV vaccines to induce particular and appropriate cellular immune responses to eliminate virus-infected cells efficiently (Cheng et al., 2018). Moreover, the E6 and E7 as oncoproteins, which are persistently encoded and expressed in HPV-infected and cancer cells, provide ideal targets for immunotherapy and the development of therapeutic vaccines against cervical cancer (Chabeda et al., 2018).

The E6 therapeutic vaccines can induce strong anti-tumor T cell-mediated immune responses, such as cytotoxic T lymphocytes, that play vital roles in the current therapeutic vaccine development. To date, numerous promising vaccine candidates for HPV-associated diseases, such as DNA and RNA vectors, viral and bacterial vectors, dendritic cells, and peptides and proteins in various combinations, have been generated and tested (Chabeda et al., 2018). However, no candidates have been approved as therapeutic vaccines for HPV-induced cancers. Among the various approaches, peptide vaccines are generally considered to be stable, safe, and easy to produce, with minimal toxicity and high specificity (Panahi, Bolhassani, Javadi, & Noormohammadi, 2018). Valuable peptide selection and interaction between peptides and human leukocyte antigen (HLA) alleles are necessary steps to induce strong immune responses. The frequency diversity of HLA genes causes binding to various peptides around the globe (Samandary et al., 2014). So, highly polymorphic HLAs must be considered to design a broad-spectrum vaccine.

In the current study, we applied two strategies to select and design peptides to produce robust responses for subsequent steps. First, to develop a board-spectrum vaccine, a *phylogenetic tree* was generated using retrieved sequences, and consensus sequences were selected.

Another strategy to increase HPV peptide-based vaccine potency is inducing CD+4 T lymphocyte cell immune responses. Although CD+8 T lymphocytes are a focal point in eliminating cancerous cells, basically CD4+ T cells assist in generating and maintaining CD+8 T cells' immune responses (de Oliveira et al., 2015). Hence, designing MHC class I and MHC class II are desirable to increase vaccine potency in pre-clinical models. In addition, the participation of both B and T cells to induce cellular and humoral immunity, respectively, is imperative to elicit a prolonged, substantial immune response.

In our study, antigenic sites of T-cells and B-cells of the E6 protein of HPV type 16 were predicted using immune computing approaches to discover candidate peptides for the development of effective peptide-based therapeutic vaccines.

2. Materials and Methods

2.1. Retrieval of the target protein

The National Center for Biotechnology Information (NCBI) database was used to retrieve the FASTA-formatted amino acid sequences of a full-length HPV16 E6 protein

(http://ncbi.nlm.nih.gov). Next, the sequences obtained were then subjected to the ClustalW2 software (T-Coffee-Server) for alignment. Bioedit software (version 7.1.3.0) was applied to trim and analyze the sequences.

2.2. Conserved region identification and phylogenetic tree

The amino acid variation in each position of the protein sequences was calculated using the Shannon entropy value. Then, the family tree was created by MEGA Software (Version.7). The maximum likelihood method was applied to draw the family tree (based on the Jones-Taylor-Thornton (JTT) model in 100 bootstrap replications). Lastly, the consensus sequence was created by selected sequences in BioEdit software.

2.3. T cell epitope prediction

For T cell epitope prediction, the IEDB MHC class I binding tool (http://tools.iedb.org/mhci/) and IEDB MHC class II binding tool (http://tools.iedb.org/mhcii/) were applied to the $CD8^+$ epitope and $CD4^+$ epitope prediction, respectively. The IEDB *recommended* method was selected, usually the consensus method as default.

2.4. Population Coverage Prediction

The IEDB Population Coverage Tool (http://tools.iedb.org/population/) was served in different regions of the world to determine the population coverage of the predicted epitopes.

2.5. Prediction of interferon gamma-inducing epitopes

The capability of the selected epitopes to induce interferon gamma IFN epitope servers was revealed using the IFN epitope server (https://crdd.osdd.net/raghava/ifnepitope/predict.php). In our study, the Support Vector Machine (SVM) algorithm and the IFN gamma versus non-IFN gamma model were used for prediction.

2.6. Linear B-Cell Epitope Prediction

One vital arm of vaccine design is the B-cell epitope, which activates B lymphocytes and stimulates a humoral immune response. Kolaskar & Tongaonkar Antigenicity (Kolaskar & Tongaonkar, 1990) was used to predict linear B-cell epitope with 75% accuracy (http://tools.iedb.org/bcell/). This method is semi-empirical and applies physicochemical characteristics of amino acid residues and their frequency of existence in experimentally known epitopes. The default threshold value (about 1.05) was considered as potential antigenic epitopes.

2.7. Molecular docking

PDB structures of both MHC (receptor) and epitope (ligand) were applied in docking. The PEP-FOLD3 online approach (https://bioserv.rpbs.univ-paris-diderot.fr/services/PEP-FOLD3/) was performed to generate 3D structures of epitopes by inputting their amino acid sequence. Also, a three-dimensional (3D) structure of HLA was found in the SWISS MODEL database (https://swissmodel.expasy.org/interactive). Hex 8.0.0 software was used to perform docking simulations with input from both MHC and peptide PDB files. Below 100 solutions, docking controls have been enabled in Shape+Electro+DARS mode. The results were reported as RSM, E-shape, and total (Bahmani, Amini-Bayat, Ranjbar, Bakhtiari, & Zarnani, 2021).

3. Results and Discussion

3.1. Protein sequences collection and primary analyses

The amino acid sequences of 1151 HPV-16 E6 proteins were recovered from the NCBI database to design a potential multi-epitope vaccine. After alignment using ClustalW software, the BioEdit software was applied to trim and analyze the aligned sequences. The calculated entropy diagram of each amino acid residue ranged from 0.0 (low variable) to 1.0 (low conservation). The position of residues with a value smaller than 0.1

(as a threshold) was defined as conserved. According to the Shannon entropy diagram (Fig. 1), HPV16-E6, a semi-conserved and excessive variation, was observed in four regions (17, 22, 85, and 90) between sequences.

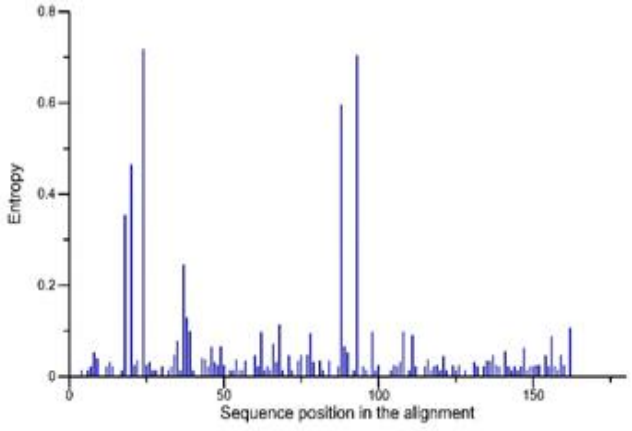

Figure. 1. Shannon's entropy plot of HPV16-E6.

The obtained result can be analyzed as a peptide engineering step to select stable and potent peptides for further excremental study. For example, in the 85th position (tyrosine comparison to histidine), the tyrosine residue could be a better choice because although it is hydrophobic, it has better interaction with MHC, provides more immunity, and also higher thermal resistance, while the histidine residue decreases the half-life of the protein. With a sufficient number of sequenced genomes, a family tree of the mutational history of a virus family can be reconstructed. So, the phylogenetic tree was constructed using the JTT model and maximum likelihood statistical method in MEGA software. At the bottom of the tree, genetic distances or the number of substitutions are indicated; the numbers at branch nodes refer to 100 bootstrap repeats (%). The mean total distance was estimated to be about 0.027. We also selected 31 isolates from different strain vaccine. For further analysis, the consensus sequence was generated at a70% sequence identity threshold. The 31 isolates and consensus sequence were generated at a 70% sequence identity threshold. The 31 isolates and consensus branches of the tree to develop a broad-spectrum sequences are summarized in (Table 1).

3.2. CD8+ and CD4+ T-Cell epitope prediction and population coverage calculation

Several bioinformatic tools have been proposed for the development of effective vaccines (Bukhari, Jain, Haq, Mehbodniya, & Webber, 2022). In 2018, Jabbar et al. used NetCTL 1.2 to predict effective MHC class I peptides (Jabbar et al., 2018). Although NetCTL 1.2 predicted the possible CD8 T cell epitopes of each protein with antigenic properties, it cannot predict the corresponding allele to these epitopes. In spite of NetCTL's capabilities, IEDB is the most widely used approach for epitope prediction, with the functionality of the IEDB extending far beyond epitope prediction. The IEDB provides tools for predicting epitopes recognized by B cells and T cells and analyzing epitope properties for more complete and reliable prediction results. This database and associated tools have been widely used in studies predicting epitopes for vaccine development, perhaps because the resource is easy to use. In this study the IEDB server was applied to predict potential peptides.

The FASTA formatted consensus sequence was applied as input for the prediction tools. The prediction was performed using the IEDB server in silico analyses tool Consensus (Kim et al., 2012), including ANN aka NetMHC (Andreatta & Nielsen, 2016; Lundegaard et al., 2008; Nielsen et al., 2003), SMM (Peters & Sette, 2005), and Comblib (Sidney et al., 2008) predictions (percentile rank <1). Among the epitopes, 41 had a smaller percentile rank indicating a high binding affinity for HLA alleles. (Table 2) presents the percentile rank of MHC I restricted alleles for the different peptide predictions.

Based on these results, a representative epitope was selected if several epitopes exhibited similar HLA binding profiles. For example, FAFRDLCIVY52-61 and AFRDLCIVYR 53-62 have binding profiles covered by FAFRDLCIVYR

Table 1. Selected HPV16 E6 Isolates from the Pedigree and Consensus Sequence

Name	**SEQUENCE**
ABC48950/E6 HPV T.16/Iran	MHQKRTAMFQDPQERPGKLPQLCTELQTTIHDIILECVYCKQQLLRREVYDFAFRDLCIV YRDGNPYAVCDKCLKFYSKISEYRHYCYSVYGTTLEQQYNKPLCDLLIRCINCQKPLCPE EKQRHLDKKQRFHNIRGRWTGRCMSCCRSSRTRRETQL
AJW82136/E6 HPV T.16/Iran	MHQKRTAMFQDPQERPRKLPQLCTELQTTIHDIILECVYCKQQLLRREVYDFAFRDLCIV YRDGNPYAVCDKCLKFYSKISEYRHYCYSLYGTTLEQQYNKPLCDLLIRCINCQKPLCPE EKQRHLDKKQRFHNIRGRWTGRCMSCCRSSRTRRETQL
AJW82151/E6 HPV T.16/Iran	MHQKRTAMFQDPQERPRKLPHLCTELQTTIHDIILECVYCKQQLLRREVYDFAFRDLCIV YRDGNPYAVCDKCLKFYSKISEYRYYCYSVYGTTLEQQYNKPLCDLLIRCINCQKPLCPE EKQRHLDKKQRFHNIRGRWTGRCMSCCRSSRTRRETQL
ALU65789/E6 HPV T.16/Brazil	MHQKRTAMFQDPQERPRKLPHLCTELQTTIHDIRLECVYCKQQLLRREVYDFAFRDLCIV YRDGNPYAVCDKCLKFYSKISEYRYYCYSVYGTTLEQQYNKPLCDLLIRCINCQKPLCPE EKQRHLDKKQRFHNIRGRWTGRCMSCCRSSRTRRETQL
ALU65767/E6 HPV T.16/Brazil	MHQKRTAMFQDPQERPRKLPHLCTELQTTIHDIILECVYCKQQLLRREVYDFAFRDLCIV YRYGNPYAVCDKCLKFYSKISEYRYYCYSVYGTTLEQQYNKPLCDLLIRCINCQKPLCPE EKQRHLDKKQRFHNIRGRWTGRCMSCCRSSRTRRETQL
ALU65786/E6 HPV T.16/Brazil	MHQKRTAMFQDPQERPRKLPQLCTELQTTIHAIILECVYCKQQLLRREVYDFAFRDLCIV YRDGNPYAVCDKCLKFYSKISEYRHYCYSLYGTTLEQQYNKPLCDLLIRCINCQKPLCPE EKQRHLDKKQRFHNIRGRWTGRCMSCCRSSRTRRETQL
AGE97044/E6 HPV T.16/Argentina	MHQKRTAMFQDPQERPRKLPDLCTELQTTIHDIILECVYCKQQLLRREVYDFAFRDLCIV YRDGNPYAVCDKCLKFYSKISEYRYYCYSLYGTTLEQQYNKPLCDLLIRCINCQKPLCPE EKQRHLDKKQRFHNIRGRWTGRCMSCCRSSRTRRETQL
ALU65779/E6 HPV T.16/Brazil	MHQKRTAMFQDPQERPRKLPQLCTELQTTIHDIILECVYCKQQLLRREVYDFAFRDLCIV YRDGNPYAVCDKCLKFYSKISEYRHYCYSLYGTTLEQQYNKPLCDLLIRCINCQKPLCPE EKQRHLDKKQRFHNIRGRWTGRCMSCCRSSRTRRETQL
ACN22512/E6 HPV T.16/Italy	MHQKRTAMFQDPQERPRKLPQLCTELQTTIHDIILECVYCNQQLLRREVYDFAFRDLCIV YRDGNPYAVCDKCLKFYSKISEYRHYCYSLYGTTLEQQYNKPLCDLLIRCINCQKPLCPE EKQRHLDKKQRFHNIRGRWTGRCMSCCRSSRTRRETQL
ACN22508/E6 HPV T.16/Italy	MHQKRTAMFQDPQERPRKLPQLCTELQTTIHDIILECVYCKQQLLRREVYDFAFRDLCIV YRDGNPYAVCDKCLKFYSKISEYRHYCYSVYGTTLEQQYNKPLCDLLIRCINCQKPLCPE EKQRHLDKKQRFHNIRGRWTGRRMSCCRSSRTRRETQL
AGS42318/E6 HPV T.16/Morocco	MHQKRTAMFQDPQERPGKLPDLCTELQTTIHDIILECVYCKQQLLRREVYDFAFRDLCIV YRDGNPYAVCEKCLKFYSKISEYRYYCYSLYGTTLEQQYNKPLCDLLIRCINCQKPLCPE EKQRHLDKKQRFHNIRGRWTGRCMSCCRSSRTRRETQL
AGS42341/E6 HPV T.16/Morocco	MHQKRTAMFQDPQERPRKLPDLCTELQTTIHDIILECVYCKQQLLRREVYDFAFRDLCIV YRDGNPYGVCDKCLKFYSKISEYRYYCYSLYGTTLEQQYNKPLCDLLIRCINCQKPLCPE EKQRHLDKKQRFHNIRGRWTGRCMSCCRSSRTRRETQL
AGS45545/E6 HPV T.16/Thailand	MHQKRTAMFQDPQERPRKLPQLCTELQTTIHDIILQCVYCKQQLLRREVYDFAFRDLCIV YRDGNPYAVCDKCLKFYSKISEYRHYCYSLYGTTLEQQYNKPLCDLLIRCINCQKPLCPE EKQRHLDKKQRFHNIRGRWTGRCMSCCRSSRTRRETQL
AFJ19771/E6 HPV T.16/Greece	MHQKRTAMFQDPQEPPRKLPQLCTELQTTIHDIILECVYCKQQLLRREVYDFAFRDLCIV YRDGNPYAVCDKCLKFYSKISEYRHYCYSVYGTTLEQQYNKPLCDLLIRCINCQKPLCPE EKQRHLDKKQRFHNIRGRWTGRCMSCCRSSRTRRETQL
AFJ19687/E6 HPV T.16/Greece	MHQKRTAMFQDPQEPPRKLPQLCTELQTTIHDIMLECVYCKQQLLRREVYDFAFRDLCIV YRDGNPYAVCDKCLKFYSKISEYRHYCYSVYGTTLEQQYNKPLCDLLIRCINCQKPLCPE EKQRHLDKKQRFHNIRGRWTGRCMSCCRSSRTCRETQL
AFJ19711/E6 HPV T.16/Greece	MHQKRTAMFQDPQEPPRKLPQLCTELQTTIHNIILECVYCKQQLLRREVYDFAFRDLCIV YRDGNPYALCDKCLKFYSKISEYRHYCYSVYGTTLEQQYNKPLCDLLIRCINCQKPLCPE EKQRHLDKKQRFHNIRGRWTGRCMSCCRLSRTRRETQL
AFJ19685/E6 HPV T.16/Greece	MHQKRTAMFQDPQEPPRKLPQLCTELQTTIHDIIIECVYCKQQLLQREVYDFAFRDLCIV YRDGNPYAVCDKCLKFYTKISEYRHYCYSVYGTTLEQQYNKPLCDLLIRCINCQKPLCPE EKQRHLDKKQRFQNIRGRWTGRCMSCCRSSRTCRETQL

Table1.Continued

AGO04505/E6 HPV T.16/Greece	MHQKRTAMFQDPQEPPRKLPQLCTELQTAIHDIILECVYCKQQLLRREVYDFAFRDLCIV YRDGNPYAVCDKCLKFYSKISEYRHYCYSLYGTTLEQQYNKPLCDLLIRCINCQKPLCPE EKQRHLDKKQRFHNIRGRWTGRCMSCCRSSRTRRETQL
AHZ30605/E6 HPV T.16/Mexico	MHQKRTAMFQDPQERPRKLPQLCTELQTTIHNIILECVYCKQQLLRREVYDFAFRDLCIV YRDGNPYAVCDKCLKFYSKISEYRHYCYSVYGTTLEQQYNKPLCDLLIRCINCQKPLCPE EKQRHLDKKQRFHNIRGRWTGRCMSCCRSSRTRRETQL
AGG35752/E6 HPV T.16/India	MHQKRTAMFQDPQERPRKLPQLCTELQTTIHDIILECVYCKQQLLRREVYDFAFRDLCIV YRDGNPYAVCDKCLKFYSKISEYRHYCYSVYGTTLEQQYNKPLCDLLIRCINCQKPLCPE EKQRHLDKKQRFHNIRGRWTGRCLSCCRSSRTRRETQL
AGS45258/E6 HPV T.16/Poland	MHQKRTAMFQDPQERPRKLPHLCTELQTTIHDIILECVYCKQQLLRREVYDFAFRDLCIV YRDGNPYAVCDKCLKFYSKISEYRYYCYSVYGTTLEQQYNKPLCDLLIRCINCQKPLCPE EKQRHLDKKQRFHNIRGRWTGRCMSCCRSSRTRRETQL
AGO04505/E6 HPV T.16/Greece	MHQKRTAMFQDPQEPPRKLPQLCTELQTAIHDIILECVYCKQQLLRREVYDFAFRDLCIV YRDGNPYAVCDKCLKFYSKISEYRHYCYSLYGTTLEQQYNKPLCDLLIRCINCQKPLCPE EKQRHLDKKQRFHNIRGRWTGRCMSCCRSSRTRRETQL
AIG19194/E6 HPV T.16/United Kingdom	MHQKRTAMFQDPQERPRKLPQLCTELQTTIHDIILECVYCKQQLLRREVYDFAFRDLCIV YRDGNPYAVCDKCLKFYSKISEYRHYCYSVYGTTLEQQYNKPLCDLLIRCINCQKPLCPE EKQRHLDKKQRFHNIRGRWTGRCMSCCRSSRTRRETQL
AGS45606/E6 HPV T.16/Nigeria	MHQKRTAMFQDPQERPGKLPDLCTELQTTIHDIILECVYCKQQLLRREVYDFAFRDLCIV YRDGNPYAVCDKCLKFYSKISEYRYYCYSLYGTTLEQQYNKPLCDLLIRCINCQKPLCPE EKQRHLDKKQRFHNIRGRWTGRCMSCCRSSRTRRETQL
AGS45343/E6 HPV T.16/Uganda	HQKRTAMFQDPQERPIKLPDLCTELQTTIHDIILECVYCKQQLLRREVYDFAFRDLCIVY RDGNPYAVCDKCLKFYSKISEYRYYCYSLYGTTLEQQYNKPLCDLLIRCINCQKPLCPEE KQRHLDKKQRFHNIRGRWTGRCMSCCRSSRTRRETQL
AAB70732/E6 HPV T.16/Uganda	MHQKRTAMFQDPQERPTKLPDLCTELQTTIHDIILECVYCKQQLLRREVYDFAFRDLCIV YRDGNPYAVCDKCLKFYSKISEYRHYCYSVYGTTLEQQYNKPLCDLLIRCINCQKPLCPD EKQRHLDKKQRFHNIRGRWTGRCMSCCRSSRTRRETQL
AGS45273/E6 HPV T.16/Paraguay	MHQKRTAMFQDPQERPRKLPHLCTELQTTIHNIILECVYCKQQLLRREVYDFAFRDLCIV YRDGNPYAVCDKCLKFYSKISEYRYYCYSVYGTTLEQQYNKPLCDLLIRCINCQKPLCPE EKQRHLDKKQRFHNIRGRWTGRCMSCCRSSRTRRETQL
AFU06625/E6 HPV T.16/Taiwan	MHQKRTAMFQDPQERPRKLPQLCTELQTTIHEIILECVYCKQQLLRREVYDFAFRDLCIV YRDGNPYAVCDKCLKFYSKISEYRHYCYSLYGTTLEQQYNKPLCDLLIRCINCQKPLCPE EKQRHLDKKQRFHNIRGRWTGRCMSCCRSSRTRRETQL
AFU06665/E6 HPV T.16/Taiwan	MHQKRTAMFQDPQERPRKLPQLCTELQTTIHDIILECVYCKQQLLRREVYDFAFRDLCIV YRDGNPYAVCDKCLKFYSKISEYRYYCYSVYGTTLEQQYNKPLCDLLIRCINCQKPLCPE EKQRHLDKKQRFHNIRGRWTGRCMSCCRSSRTRRETQL
AHZ96690/E6 HPV T.16/Spain	MHQKRTAMFQDPQERPTKLPDLCTELQTTIHDIILECVYCKQQLLRREVYDFAFRDLCIV YRDGNPYAVCDKCLKFYSKISEYRYYCYSLYGTTLEQQYNKPLCDLLIRCINCQKPLCPE EKQRHLDKKQRFHNIRGRWTGRCMSCCRSSRTRRETQL
AHN92524/E6 HPV T.16/Spain	MHQKRTAMFQDPQERPRKLPQLCTELQTTIHDILSECVYCKQQLLRREVYDFAFRDLCTV YRDGNPYAVCDKCLQFYSKISEYRHYCYSVYGTTSEQQYNKPLCDLLLRCINCQKPLCPE EKQRHLDKKQRFHNIRGRWTGRCMSCCRSSRTRRETQL
AFP44447/E6 HPV T.16/Costa Rica	MHQKRTAMFQDPQERPRKLPQLCTELQTTIYDIILECVYCKQQLLRREVYDFAFRDLCIV YRDGNPYAVCDKCLKFYSKISEYRHYCYSLYGTTLEQQYNKPLCDLLIRCINCQKPLCPE EKQRHLDKKQRFHNIRGRWTGRCMSCCRSSRTRRETQL
Consensus	MHQKRTAMFQDPQERPRKLPQLCTELQTTIHAIILECVYCKQQLLRREVYDFAFRDLCIV YRDGNPYAVCDKCLKFYSKISEYRHYCYSLYGTTLEQQYNKPLCDLLIRCINCQKPLCPE EKQRHLDKKQRFHNIRGRWTGRCMSCCRSSRTRRETQL

FAFRDLCIVYR 52-62. Using this approach, the epitopes were reduced from 41 to 17 (Table 3). In addition, the IEDB-PPC tool was used to determine the population coverage (PC) of the 17 selected MHC class I. The distribution of MHC alleles varies in different geographic regions and ethnic groups around the world. In this study, these peptides and their alleles cover 75.87% of

the world population. The highest population coverage is related to KISEYRHYCYSL (36.35%) followed by IHAIILECVYCK, with a population coverage of 32.38%. The lowest coverage was found for WTGRCMSCCRSSR (7.09%). According to population coverage, six peptides were selected for molecular docking study: KISEYRHYCYSL, IHAIILECVYCK, GTTLEQQYNKPL, CIVYRDGNPYA, CYSLYGTTL, and MHQKRTAMF.

Table 2. Binding Profile of the Conserved MHC I Epitopes of HPV16 E6

Epitope Sequences	MHC I-restricted alleles	Percentile Rank*
FAFRDLCIVY**	HLA-B*35:01	0.09
	HLA-B*53:01	0.85
	HLA-B*51:01	0.4
	HLA-A*02:06	0.6
AFRDLCIVYR	HLA-A*31:01/0.16	0.16
	HLA-A*33:01/0.23	0.23
KFYSKISEYR	HLA-A*31:01	0.11
	HLA-A*68:01	0.26
	HLA-A*33:01	0.27
	HLA-A*30:02	0.3
	HLA-A*33:01	0.47
	HLA-A*31:010	0.87
RCMSCCRSSR	HLA-A*31:01	0.16
	HLA-A*33:01	0.17
	HLA-A*31:01	0.51
MSCCRSSRTR	HLA-A*33:01	0.41
	HLA-A*31:01	0.47
WTGRCMSCCR	HLA-A*33:01	0.68
KISEYRHYCYSL	HLA-A*01:01	0.2
	HLA-B*40:01	0.26
	HLA-B*44:03	0.4
	HLA-A*30:02	0.47
	HLA-A*01:01	0.65
	HLA-A*30:02	0.83
	HLA-B*44:02	0.87
IHAIILECVYCK	HLA-B*35:01	0.2
	HLA-A*30:02	0.81
	HLA-B*15:01	0.95
	HLA-A*30:02	0.99
IILECVYCK	HLA-A*11:01	0.44
	HLA-A*11:01	0.49
	HLA-A*03:01	0.96
GTTLEQQYNKPL	HLA-A*11:01	0.22
	HLA-B*40:01	0.38
	HLA-A*11:01	0.81
	HLA-A*68:01	1.00
KKQRFHNIRGR	HLA-A*31:01	0.23
	HLA-A*30:01	0.3
	HLA-A*31:01	0.91
DKKQRFHNI	HLA-B*08:01	0.3
YAVCDKCLKF	HLA-B*53:01	0.26
	HLA-B*35:01	0.8
	HLA-A*68:01	0.9
NIRGRWTGRCM	HLA-A*33:01	0.31
	HLA-B*07:02	0.6
KQQLLRREVY	HLA-B*15:01	0.36
TAMFQDPQERPR	HLA-A*33:01	0.38
	HLA-A*68:01	0.43
IVYRDGNPYA	HLA-A*30:01	0.4
	HLA-A*30:01	0.42
	HLA-B*15:01	0.7
	HLA-A*30:02	0.77
	HLA-B*35:01	0.9
CIVYRDGNPY	HLA-A*26:01	0.54
LLIRCINCQK	HLA-A*30:01	0.4
	HLA-A*03:01	0.76
CVYCKQQLLR	HLA-A*68:01	0.55
	HLA-A*33:01	0.86
	HLA-A*33:01	0.91
CYSLYGTTL	HLA-A*24:02	0.58
KQRHLDKKQR	HLA-A*31:01	0.9
MHQKRTAMF	HLA-A*24:02	0.97

Table2.Continued

All analyses were performed by the Consensus method.
** The italic, bold, and underlined epitopes are associated with italic, bold, and underlined cases in the MHC-I restricted alleles/percentile rank and methods used columns, respectively.

Table 3. Binding Profile and Population Protection Coverage of the Selected Conserved Epitopes

	MHC I-restricted alleles	**Population coverage (%) (world)**
FAFRDLCIVYR	HLA-B*35:01,HLA-B*53:01,HLA-B*51:01,HLA-A*02:06, HLA-A*31:01, HLA-A*33:01	25.22
KFYSKISEYR	HLA-A*31:01,HLA-A*68:01,HLA-A*33:01,HLA-A*30:02	14.98
WTGRCMSCCRSSR	HLA-A*31:01, HLA-A*33:01	7.09
KISEYRHYCYSL	HLA-A*01:01,HLA-B*40:01, HLA-B*44:03, HLA-A*30:02, HLA-A*01:01, HLA-A*30:02, HLA-B*44:02	36.35
IHAIILECVYCK	HLA-B*35:01, HLA-A*30:02, HLA-B*15:01, HLA-A*30:02, HLA-A*11:01, HLA-A*03:01	32.38
GTTLEQQYNKPL	HLA-A*11:01, HLA-B*40:01, HLA-A*68:01	27.06
DKKQRFHNIRGR	HLA-A*31:01,HLA-A*30:01 ,HLA-B*08:01	18.73
YAVCDKCLKF	HLA-B*53:01, HLA-B*35:01, HLA-A*68:01	16.04
NIRGRWTGRCM	HLA-A*33:01, HLA-B*07:02	14.32
KQQLLRREVY	HLA-B*15:01	8.44
TAMFQDPQERPR	HLA-A*33:01,HLA-A*68:01	7.56
CIVYRDGNPYA	HLA-A*30:01, HLA-B*15:01, HLA-A*30:02, HLA-B*35:01 HLA-A*26:01	26.42
LLIRCINCQK	HLA-A*30:01, HLA-A*03:01	20.35
CVYCKQQLLR	HLA-A*68:01, HLA-A*33:01, HLA-A*33:01	7.56
CYSLYGTTL	HLA-A*24:02	21.38
KQRHLDKKQR	HLA-A*31:01	5.36

As stated above, the IEDB predictor was used based on MHC-II (HLA-DR and DQ) to generate epitope-HLA pairs. By considering percentile rank (>6) and the *Consensus* tool (Wang et al., 2008; Wang et al., 2010), 19 peptides were discovered with a high binding affinity for HLA alleles. (Table 4) shows the percentile rank, MHC-II restricted alleles for different peptide predictions, and population coverage. According to population coverage, three peptides were selected for the molecular docking study:
QQLLRREVYDFAFRDL,
KQQLLRREVYDFAFRD,
QLLRREVYDFAFRDLC.

Six MHC class I and three MHC class II epitopes were combined to make potential global vaccines with population coverage of 86.41%.

Table 4. Binding Profile of the Conserved MHC-II Epitopes

Epitope Sequences	MHC II-restricted alleles/ Percentile Rank*	Population coverage (%)(world)
EYRHYCYSLYGTTLEQ	HLA-DRB1*04:05/3.10	3.02
SEYRHYCYSLYGTTLE	HLA-DRB1*04:05/3.10	3.02
YRHYCYSLYGTTLEQQ	HLA-DRB1*04:05/3.10	3.02
RHYCYSLYGTTLEQQY	HLA-DRB1*04:05/3.90	3.02
HYCYSLYGTTLEQQYN	HLA-DRB1*04:05/4.40	3.02
QQLLRREVYDFAFRDL	HLA-DRB3*01:01/3.20	43.67
	HLA-DPA1*02:01/DPB1*05:01/4.25	
KQQLLRREVYDFAFRD	HLA-DRB3*01:01/3.40	-
LRREVYDFAFRDLCIV	HLA-DPA1*02:01/DPB1*05:01/3.80	43.67
QLLRREVYDFAFRDLC	HLA-DRB3*01:01/3.80	43.67
	HLA-DPA1*02:01/DPB1*05:01/4.25	
LLRREVYDFAFRDLCI	HLA-DPA1*02:01/DPB1*05:01/4.10	43.67
	HLA-DRB3*01:01/4.80	
KCLKFYSKISEYRHYC	HLA-DRB5*01:01/4.90	-
CLKFYSKISEYRHYCY	HLA-DRB5*01:01/5.00	-
DKCLKFYSKISEYRHY	HLADRB1*15:01/5.10, HLADRB5*01:01/5.10	18.41
LKFYSKISEYRHYCYS	HLA-DRB5*01:01/5.20	-
CTELQTTIHAIILECV	HLA-DQA1*01:02/DQB1*06:02/5.05	34.55
ELQTTIHAIILECVYC	HLA-DQA1*01:02/DQB1*06:02/5.10	34.55
TELQTTIHAIILECVY	HLA-DQA1*01:02/DQB1*06:02/5.10	34.55
LQTTIHAIILECVYCK	HLA-DQA1*01:02/DQB1*06:02/5.15	34.55
QTTIHAIILECVYCKQ	HLA-DQA1*01:02/DQB1*06:02/5.20	34.55

* All analyses were performed by the Consensus method

Predicting target epitopes via immune computing techniques leads to cost-effective enhancement of T cell immune response (Melief, 2018). In our study, immune computing approaches were used to predict and assess the binding affinity of candidate E6 protein peptides (from HPV types 16). Taking into account percentile rank, population coverage, and global energy, six MHC class I epitopes, KISEYRHYCYSL, IHAIILECVYCK, GTTLEQQYNKPL, CIVYRDGNPYA, CYSLYGTTL, MHQKRTAMF, and three MHC class II epitopes, QQLLRREVYDFAFRDL, KQQLLRREVYDFAFRD,and QLLRREVYDFAFRDLC, were selected.

Jabbar et al. predicted the E6 and E7 antigenic peptides of HPV16 and HPV18 using immunoinformatic methods and then analyzed their ability to bind MHC-I molecules using

molecular docking and MD simulation (2018). Yao et al. (2013) predicted 59 and 22 CTL epitopes of E6 and E7, respectively, using the IEDB server for the most common alleles in the world population. In other studies, only chimeric structures containing the epitopes of E5, E6, and E7 proteins were designed, and no further bioinformatics evaluations were performed on the structure. So far, only one study is similar to our investigation; they developed a multi-epitope chimeric vaccine using the most oncogenic strains, HPV 16 and HPV 18, through an immune computing approach. They used the L1, E5, E6, and E7 oncoproteins from both HPV 16 and HPV 18 strains for epitope prediction. The recombinant chimeric vaccine construct consists of selected helper and cytotoxic T cell epitopes (Kumar, Sahu, Kumari, Dixit, & Khare, 2022). In contrast to previous studies that used only one sequence to predict peptides, in the present study, a *phylogenetic tree* was generated using the retrieved sequences, and the consensus sequences were selected to develop a board-spectrum vaccine.

KISEYRHYCYSL binds with high affinity to HLA-A*01:01, HLA-B*40:01, HLA-B*44:03, HLA-A*30:02, HLA-A*01:01, HLA-A*30:02, HLA-B*44:02. It has high PC in Central Africa (24.15%), East Africa (29.64%), East Asia (23.13%), Europe 45.83%, North Africa (31.13%), North America (37.58%), Northeast Asia (20.23%), Oceania (26.69%), South Africa (36.71%), South Africa (36.71%), South Asia (21.68%), Southeast Asia (29.44%), Southwest Asia (24.22%), West Africa (24.15%), and West Indies (41.79%).

The 2nd selected peptide, IHAIILECVYCK, binds with high affinity for HLA-B*35:01, HLA-A*30:02, HLA-B*15:01, HLA-A*30:02, HLA-A*11:01, HLA-A*03:01 with great PC in Central Africa (33.61%), North Africa (29.53%), North America (42.07%), Northeast Asia (56.07%), South Africa (35.32%), Southeast Asia (41.44%), Southwest Asia (29.73%), West Africa (40.41%), and West Indies (40.83%).

The GTTLEQQYNKPL, the 3rd selected peptide, binds to HLA-A*11:01, HLA-B*40:01, and HLA-A*68:01 with high PC in North America (25.61%), Northeast Asia (52.76%), South America (20.01%), South Asia (33.64%), and Southeast Asia (52.26%).

The CIVYRDGNPYA, the 4th selected peptide, has an affinity for HLA-A*30:01, HLA-B*15:01, HLA-A*30:02, HLA-B*35:01, and HLA-A*26:01 molecules and has high PC in Central Africa (28.96%), East Africa (30.75%), East Asia (39.71%), Europe (29.59%), North Africa (30.85%), North America (32.46%), Northeast Asia (24.47%), South Africa (30.67%), South Asia (25.65%), Southwest Asia (22.59%), West Africa (40.99%), and West Indies (29.08%).

The 5th selected peptide, CYSLYGTTL, binds with high affinity for HLA-A*24:02 with high PC in East Asia (49.65%), North America (22.87%), Northeast Asia (24.09%), Oceania (52.33%), South America (23.76%), and Southeast Asia (40.08%).

The MHQKRTAMF, the 6th peptide, has a great affinity to HLA-A*24:02 with high PC in East Asia (49.65%), North America (22.87%), Northeast Asia (24.09%), Oceania (52.33%), South America (23.76%), and Southeast Asia (40.08%).

The QQLLRREVYDFAFRDL and QLLRREVYDFAFRDLC, MHC-II peptides, bind to HLA-DRB3*01:01 and HLA-DPA1*02:01/DPB1*05:01 with great PC in Central Africa (37.59%), East Africa (36.62%), East Asia (59.33%), Europe (23.63%), Northeast Asia (56.6%), Oceania (70.68%), South America (40.21%), South Asia (39.01%), Southeast Asia (55.37%), and West Africa (71.67%).

A combination of selected epitopes creates a potential global vaccine with 86.41% population

protection coverage. The greatest number of cases of cervical cancer (285,000; 54.0%) and expiration (144,000; 54.7%) are found in Asia, followed by Africa with 99.000 new cases and 60.000 expiration and the Americas (83,000 new cases and 36,000 expiration). The substantial rate in China and India is 62,000 and 123,000 new patients and 30,000 and 67,000 expirations, respectively (Serrano et al. 2017). So, multiple epitope combinations will be valuable in Central Africa (69.72%), China (90.68%), East Africa (69.92%), East Asia (92.55%), India (81.11%), North Africa (64.74%), Northeast Asia (90.97%), South Africa (69.98%), South Asia (84.55%), Southeast Asia (91.26%), Southwest Asia (60.36%), and West Africa (88.9%).

Some of the selected peptides were tested and validated experimentally. For example, a 1999 study found that VYRDGNPYA linked to minimal B-epitopes could induce T-helpers for the production of cognate antibodies (Azoury-Ziadeh, Herd, Fernando, Frazer, & Tindle, 1999). In 2000, Bourgault Villada et al. determined that ISEYRHYCY related to HLA-B18 could be identified by CD8+ T cells in healthy donors and patients with cervical intraepithelial neoplasia grade 3 (Bourgault Villada et al., 2010). Other studies showed that IILECVYCK (Kast et al., 1994), TTLEQQYNK (Kast et al., 1994), GTTLEQQYNK (Kast et al., 1994), IVYRDGNPY (Bourgault Villada et al., 2010; Kast et al., 1994), VYRDGNPYA (Liao et al., 2015), CYSLYGTTL (Kast et al., 1994; Mizuuchi et al., 2012), and MHQKRTAMF (Kast et al., 1994) were reported to bind to HLAs with high affinity.

3.3. Prediction IFN-γ-inducing epitopes

The production of IFN-γ is a vital part of the immune system to control cancer and infection, in this study, selected E6 peptides were analyzed by the IFN epitope server. As shown in (Table 5), CYSLYGTTL can induce the generation of IFN-γ

Table 5. IFN- γ Inducing Scores of the Engineered E6 Epitopes

Epitope sequence	Score
KISEYRHYCYSL	-1.1267727
IHAIILECVYCK	-0.31585139
GTTLEQQYNKPL	-0.9312477
CIVYRDGNPYA	-1.0783295
CYSLYGTTL	0.32817293(POSITIVE)
MHQKRTAMF	-0.57038234
QQLLRREVYDFAFRDL	-0.63716102
KQQLLRREVYDFAFRD	-0.34931085
QLLRREVYDFAFRDLC	-0.6535803

3.4. Linear B-Cell Epitope Prediction

The Kolaskar and Tongaonkar method was applied to predict linear B-cell epitopes using protein structures such as hydrophobicity and charge (e.g., Cys, Leu, and Val) on a protein's surface. Taking the threshold into account, seven linear peptides were predicted as HPV16 E6 B-cell linear epitopes (consensus sequence), see (Table 6). The graphical representation of the predicted epitopes is shown in (Fig 2).

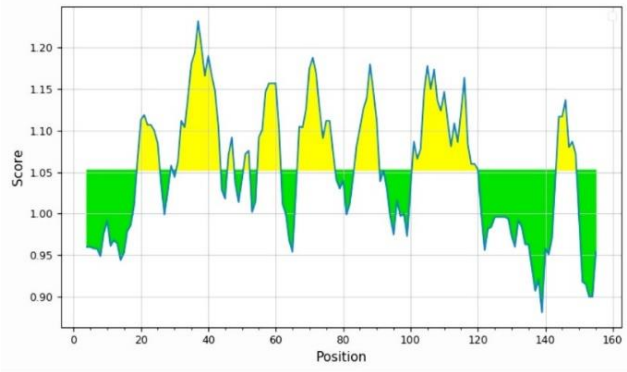

Figure. 2. B-cell epitope prediction based on the Kolaskar and Tongaonkar antigenicity scale with a threshold of 1.06.

Table 6. Linear B Cell Epitopes of HPV16-E6

Peptide*	Start	End	Length
LPQLCTE	19	25	7
HAIILECVYCKQQ	31	43	13
RDLCIVY	55	61	7
YAVCDKCLKFY	67	77	11
RHYCYSL	84	90	7
KPLCDLLIRCINCQKPLCPE	101	120	20
MSCCRS	144	149	6

*Bold residues were also predicted to be CTL epitopes

Linear epitopes have a more stable character than conformational epitopes, so for B-cell epitopes prediction, only linear epitopes were predicted (Bahmani, Amini-Bayat, Ranjbar, Bakhtiari, & Zarnani, 2021). So, among the selected peptides, five can induce CTL responses and antibodies as humoral responses, and one can stimulate IFN-γ secretion.

3.5. Docking of T-cell Epitopes to MHC Molecules

For the molecular study, six selected MHC-I epitopes and three selected MHC-II epitopes were analyzed using Hex 8.0.0 software. Each epitope was paired with an appropriate HLA molecule that is estimated to bind firmly and liberate more binding energy. CD8+ MHC class I epitope interactions with various HLA alleles in the docking assay are shown in (Table 7). In addition, the 27-docking complex was performed and displayed a variety of binding affinities indicated as global energy (-403.03 to -645.98 kcal/mol), RSM-1.

(Table 8) presents the CD4+ MHC class II epitopes with different HLA alleles in a docking study with a range of binding affinities in relation to the global energy (-329.73 to 696.22 Kcal/mol), RSM −1. Peptide-docked poses are shown in (Fig. 3)

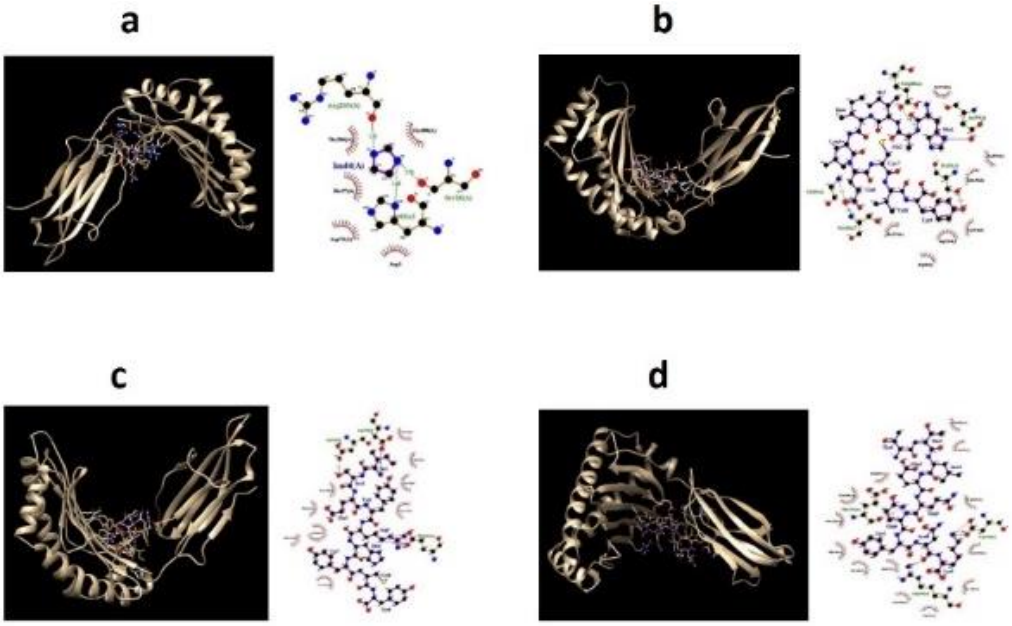

Figure. 3. Selected epitopes at their MHC-I binding site. The MHQKRTAMF, HAIILECVY, ISEYRHYCY, and TTLEQQYNK were randomly selected from high global energy. (a) Peptide MHQKRTAMF at its receptor HLA-A*24:02. (b) Peptide HAIILECVY at its receptor HLA-B*35:01. (c) Peptide ISEYRHYCY at its receptor HLA-B*44:02. (d) Peptide TTLEQQYNK at its receptor HLA-A*68:01

Table 7. Energetics of HPV16 E6 Peptide-MHC-I Complexes

HLA/Peptide	H-bond	Bumps	RSM	Total	E shape	E force	E air
1.ISEYRHYCY (80-88)							
HLA-A*01:01	-1	-1	-1.00	-490.16	-490.16	0.00	0.00
HLA-A*30:02	-1	-1	-1.00	-492.37	-492.37	0.00	0.00
HLA-B*44:02	-1	-1	-1.00	-477.80	-477.80	0.00	0.00
2.SEYRHYCYSL	(81-90)						
HLA-B*40:01	-1	-1	-1.00	-484.15	-484.15	0.00	0.00
HLA-B*44:02	-1	-1	-1.00	-512.16	-512.16	0.00	0.00
HLA-B*44:03	-1	-1	-1.00	-505.20	-505.20	0.00	0.00
3.KISEYRHYCY	(79-88)						
HLA-A*01:01	-1	-1	-1.00	-567.62	-567.62	0.00	0.00
HLA-A*30:02	-1	-1	-1.00	-575.80	-575.80	0.00	0.00
4.HAIILECVY (31-	39)						
HLA-A*30:02	-1	-1	-1.00	-421.00	-421.00	0.00	0.00
HLA-B*35:01	-1	-1	-1.00	-403.03	-403.03	0.00	0.00
5.IHAIILECVY(30-	39)						
HLA-A*30:02	-1	-1	-1.00	-409.98	-409.98	0.00	0.00
HLA-B*15:01	-1	-1	-1.00	-438.80	-438.80	0.00	0.00
6.AIILECVYCK (32-41)							
HLA-A*11:01	-1	-1	-1.00	-467.95	-467.95	0.00	0.00
7.IILECVYCK (33-41)							
HLA-A*03:01	-1	-1	-1.00	-457.08	-457.08	0.00	0.00
HLA-A*11:01	-1	-1	-1.00	-513.29	-513.29	0.00	0.00
8.TTLEQQYNK (93-101)							
HLA-A*11:01	-1	-1	-1.00	-465.07	-465.07	0.00	0.00
HLA-A*68:01	-1	-1	-1.00	-435.54	-435.54	0.00	0.00
9.LEQQYNKPL(95-103)							
HLA-B*40:01	-1	-1	-1.00	-470.76	-470.76	0.00	0.00
10.GTTLEQQYNK (92-101)							
HLA-A*11:01	-1	-1	-1.00	-489.63	-489.63	0.00	0.00
11.IVYRDGNPY (59-67)							
HLA-A*30:02	-1	-1	-1.00	-459.03	-459.03	0.00	0.00
HLA-B*15:01	-1	-1	-1.00	-467.54	-467.54	0.00	0.00
HLA-B*35:01	-1	-1	-1.00	-469.78	-469.78	0.00	0.00
12.VYRDGNPYA (60-68)							
HLA-A*30:01	-1	-1	-1.00	-497.74	-497.74	0.00	0.00
13.IVYRDGNPYA (59-68)							
HLA-A*30:01	-1	-1	-1.00	-452.59	-452.59	0.00	0.00
14. CIVYRDGNPY (58-67)							
HLA-A*26:01	-1	-1	-1.00	-460.92	-460.92	0.00	0.00
15. CYSLYGTTL (87-96)							
HLA-A*24:02	-1	-1	-1.00	-473.53	-473.53	0.00	0.00
16. MHQKRTAMF (1-9)							
HLA-A*24:02	-1	-1	-1.00	-645.98	-645.98	0.00	0.00

*Corresponding author. Tel: *+982157416815*
E-mail address: Amini-bayat@irost.org, Zahra_aminibaiat@yahoo.com.
DOI: 10.22104/ARMMT.2022.1232

Table 8. Energetics of the HPV16-E6 Peptide-MHC-II Complexes

HLA/Peptide	H-bond	Bumps	RSM	Total	E shape	E force	E air
1.QQLLRREVYDFAFRDL (42-57)							
HLA-DRB3*01:01	-1	-1	-1	-329.73	-329.73	0.00	0.00
HLA-DPA1*02:01	-1	-1	-1	-696.22	-696.22	0.00	0.00
HLA-DPB1*05:01	-1	-1	-1	-484.00	-484.00	0.00	0.00
2.KQQLLRREVYDFAFRD(41-56)							
HLA-DRB3*01:01	-1	-1	-1	-357.23	-357.23	0.00	0.00
3.QLLRREVYDFAFRDLC (43-58)							
HLA-DRB3*01:01	-1	-1	-1	-343.30	-343.30	0.00	0.00
HLA-DPA1*02:01	-1	-1	-1	-672.56	-672.56	0.00	0.00
HLA-DPB1*05:01	-1	-1	-1	-509.40	-509.40	0.00	0.00

The binding affinity of the predicted peptide fragments was probed in the peptide-MHC interactions in molecular docking studies. There are several docking methods to calculate the global energy of peptide-MHC interactions. However, due to limitations, they cannot always create docking poses that resemble the experimentally verified binding modes (Jabbar et al., 2018). Therefore, the HEX 8.0.0 was proposed because of its advantages, such as B. independent of the structure and binding site of the peptide and the preference of the spherical polar Fourier (SPF) correlation algorithm over the fast Fourier transform (FFT).

4. Conclusion

In conclusion, bioinformatic approaches and in silico analysis were used to predict capable antigenic peptides to develop therapeutic HPV vaccines. Our predicted peptides could serve as a therapeutic HPV vaccine with a population coverage of 86.41%. These theoretically

confirmed peptides could be used in a polyepitope construct as vaccine candidates for further analysis. In addition, these results provide new insights into the development of therapeutic vaccines.

As explained at the beginning of the results, unlike E7, the E6 oncoprotein is a protein that has undergone many mutations and changes in different sequences reported from different areas in the world. For this reason, instead of using a single sequence, all sequences should be

collected to create a broad-spectrum vaccine. Unlike other studies, this was achieved in this study, and the phylogeny tree was drawn and subsequent analyzes were performed on the consensus sequence. It should be noted that the results of this research *in vitro* and *in vivo* studies have achieved significant results that will be published soon.

Abbreviation

HPV: Human papillomaviruses; MHC: Major histocompatibility complex; HLA: Leukocyte antigen; PC: Population coverage.

Conflict of Interest

The authors declare that they have no competing interests.

Acknowledgments

This work was supported financially by Cancer Control Research Center, Cancer Control Foundation, Iran University of Medical Sciences, Tehran, Iran, as project No: CCF-97087. Also, we are thankful for supporting the project by Iranian research organization for science and technology (IROST).

Ethical approval

This article does/does not contain any studies with human participants or animals performed by any of the authors. This article does/does not contain any studies with human participants or animals performed by any of the authors.

Open access

References

[1]Andreatta, M., & Nielsen, M. (2016). Gapped sequence alignment using artificial neural networks: application to the MHC class I system. Bioinformatics, 32(4), 511-517.

[2]Azoury-Ziadeh, R., Herd, K., Fernando, G. J., Frazer, I. H., & Tindle, R. W. (1999). T-helper epitopes identified within the E6 transforming protein of cervical cancer-associated human papillomavirus type 16. Viral immunology, 12(4), 297-312.

[3]Bahmani, B., Amini-Bayat, Z., Ranjbar, M. M., Bakhtiari, N., & Zarnani, A.-H. (2021). HPV16-E7 protein T cell epitope prediction and global therapeutic peptide vaccine design based on human leukocyte antigen frequency: an in-silico study. International Journal of Peptide Research and Therapeutics, 27(1), 365-378.

[4]Bourgault Villada, I., Moyal Barracco, M., Berville, S., Bafounta, M., Longvert, C., Premel, V., . . . Paniel, B. (2010). Human papillomavirus 16-specific T cell responses in classic HPV-related vulvar intra-epithelial neoplasia. Determination of strongly immunogenic regions from E6 and E7 proteins. Clinical & Experimental Immunology, 159(1), 45-56.

[5]Bukhari, S. N. H., Jain, A., Haq, E., Mehbodniya, A., & Webber, J. (2022). Machine learning techniques for the prediction of B-cell and T-cell epitopes as potential vaccine targets with a specific focus on SARS-CoV-2 pathogen: A review. Pathogens, 11(2), 146.

[6]Jabbar, B., Rafique, S., Salo-Ahen, O. M., Ali, A., Munir, M., Idrees, M., . . . Jabbar, I. (2018). Antigenic peptide prediction from E6 and E7 oncoproteins of HPV types 16 and 18 for therapeutic vaccine design using immunoinformatics and MD simulation analysis. Frontiers in immunology, 9, 3000.

[7]Kast, W. M., Brandt, R., Sidney, J., Drijfhout, J.-W., Kubo, R. T., Grey, H. M., . . . Sette, A. (1994). Role of HLA-A motifs in identification of potential CTL epitopes in human papillomavirus type 16 E6 and E7 proteins. The Journal of Immunology, 152(8), 3904-3912.

[8]Kim, Y., Ponomarenko, J., Zhu, Z., Tamang, D., Wang, P., Greenbaum, J., . . . Bourne, P. E. (2012). Immune epitope database analysis resource. Nucleic acids research, 40(W1), W525-W530.

[9]Kumar, A., Sahu, U., Kumari, P., Dixit, A., & Khare, P. (2022). Designing of multi-epitope chimeric vaccine using immunoinformatic platform by targeting oncogenic strain HPV 16 and 18 against cervical cancer. Scientific Reports, 12(1), 1-16.

[10]Liao, S., Zhang, W., Hu, X., Wang, W., Deng, D., Wang, H., . . . Zhang, H. (2015). A novel "priming-boosting" strategy for immune interventions in cervical cancer. Molecular immunology, 64(2), 295-305.

[11]Lundegaard, C., Lamberth, K., Harndahl, M., Buus, S., Lund, O., & Nielsen, M. (2008). NetMHC-3.0: accurate web accessible predictions of human, mouse and monkey MHC class I affinities for peptides of length 8–11. Nucleic acids research, 36(suppl_2), W509-W512.

[12]Mizuuchi, M., Hirohashi, Y., Torigoe, T., Kuroda, T., Yasuda, K., Shimizu, Y., . . . Sato, N. (2012). Novel oligomannose liposome-DNA complex DNA vaccination efficiently evokes anti-HPV E6 and E7 CTL responses. Experimental and molecular pathology, 92(1), 185-190.

[13]Nielsen, M., Lundegaard, C., Worning, P., Lauemøller, S. L., Lamberth, K., Buus, S., . . . Lund, O. (2003). Reliable prediction of T-cell epitopes using neural networks with novel sequence representations. Protein Science, 12(5), 1007-1017.

[14]Peters, B., & Sette, A. (2005). Generating quantitative models describing the sequence specificity of biological processes with the stabilized matrix method. BMC bioinformatics, 6(1), 132.

[15]Sidney, J., Assarsson, E., Moore, C., Ngo, S., Pinilla, C., Sette, A., & Peters, B. (2008). Quantitative peptide binding motifs for 19 human and mouse MHC class I

molecules derived using positional scanning combinatorial peptide libraries. Immunome research, 4(1), 2.

[16]Wang, P., Sidney, J., Dow, C., Mothé, B., Sette, A., & Peters, B. (2008). A systematic assessment of MHC class II peptide binding predictions and evaluation of a consensus approach. PLoS computational biology, 4(4).

[17]Wang, P., Sidney, J., Kim, Y., Sette, A., Lund, O., Nielsen, M., & Peters, B. (2010). Peptide binding predictions for HLA DR, DP and DQ molecules. BMC bioinformatics, 11(1), 568.

[18]Yao, Y., Huang, W., Yang, X., Sun, W., Liu, X., Cun, W., & Ma, Y. (2013). HPV-16 E6 and E7 protein T cell epitopes prediction analysis based on distributions of HLA-A loci across populations: an in silico approach. Vaccine, 31(18), 2289-2294.