Infection, Genetics and Evolution 95 (2021) 105084

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Research papers

Designing a therapeutic and prophylactic candidate vaccine against human papillomavirus through vaccinomics approaches

Ashkan Bagheri [a,b], Navid Nezafat [a], Mahboobeh Eslami [a], Younes Ghasemi [a,b], Manica Negahdaripour [a,b,*]

[a] Pharmaceutical Science Research Center, Shiraz University of Medical Science, Shiraz, Iran
[b] Department of Pharmaceutical Biotechnology, School of Pharmacy, Shiraz University of Medical Sciences, Shiraz, Iran

ARTICLE INFO

Keywords:
In silico
Multi-epitope vaccine
Subunit vaccine
Cervical cancer (CxCa)
HPV
Vaccine design

ABSTRACT

Objective: Human papillomavirus (HPV) is the main cause of cervical cancer, the 4th prominent cause of death in women globally. Previous vaccine development projects have led to several approved prophylactic vaccines available commercially, all of which are made using major capsid-based (L1). Administration of minor capsid protein (L2) gave rise to the second generation investigational prophylactic HPV vaccines, none of which are approved yet due to low immunogenicity provided by the L2 capsid protein. On the other hand, post-translation proteins, E6 and E7, have been utilized to develop experimental therapeutic vaccines. Here, *in silico* designing of a therapeutic and prophylactic vaccine against HPV16 is performed.

Methods: In this study, several immunoinformatic and computational tools were administered to identify and design a vaccine construct with dual prophylactic and therapeutic applications consisting of several epitope regions on L2, E6, and E7 proteins of HPV16.

Results: Immunodominant epitope regions (aa 12–23 and 78–78 of L2 protein, aa 11–27 of E6 protein, and aa 70–89 of E7 protein) were employed, which offered adequate immunogenicity to induce immune responses. Resuscitation-promoting factors (RpfB and RpfE) of *Mycobacterium tuberculosis* were integrated in two separate constructs as TLR4 agonists to act as vaccine adjuvants. Following physiochemical and structural evaluations carried out by various bioinformatics tools, the designed constructs were modeled and validated, resulting in two 3D structures. Molecular docking and molecular dynamic simulations suggested stable ligand-receptor interactions between the designed construct and TLR4.

Conclusion: Ultimately, this study led to suggest the designed construct as a potential vaccine candidate with both prophylactic and therapeutic applications against HPV by promoting Th1, Th2, CTL, and B cell immune responses, which should be further confirmed in experimental studies.

1. Introduction

Cervical cancer is the 4th prominent cause of death in women globally (Torre et al., 2015), which is mainly caused by human papillomaviruses (HPVs). HPV is one of the most common sexually transmitted diseases worldwide. HPVs, a group of DNA viruses, are known as one of the most important causes of sexually transmitted diseases. Currently, 80 types of HPV have been identified including HPV16 and HPV-18 (Severson et al., 2001). Nearly 99% of cervical cancers are due to HPVs of which 70% are caused by HPV16 and HPV-18 (Bruni et al., 2010).

Considering the importance of HPV infections, different preventive strategies such as designing prophylactic vaccines are employed to combat it (Harper and Demars, 2014). Initially, prophylactic HPV vaccines were developed by using the major capsid protein of HPV, called L1 protein, which mainly activates the humoral pathway. The three currently available HPV vaccines in global market, Cervarix®, Gardasil®, and Gardasil 9®, are prophylactic vaccines containing L1 capsid protein (Panatto et al., 2015).

Another important minor capsid protein of HPV is L2, which is a short linear protein. L2 provokes the immunologic responses with the same mechanism as L1. However in contrast to L1, the L2 capsid protein

* Corresponding author at: Department of Pharmaceutical Biotechnology, School of Pharmacy, Shiraz University of Medical Sciences, P.O. Box 71345-1583, Shiraz, Iran.
E-mail address: negahdaripour@sums.ac.ir (M. Negahdaripour).

https://doi.org/10.1016/j.meegid.2021.105084
Received 12 May 2021; Received in revised form 6 September 2021; Accepted 11 September 2021
Available online 20 September 2021
1567-1348/© 2021 Elsevier B.V. All rights reserved.

has a well-conserved sequence in different types of HPV thereby providing a wider protection (Wang et al., 2015). Of note, only the N-terminus domain of L2 is exposed during the infection cycle (Campo and Roden, 2010).

However, L2 capsid protein usage causes low immunologic responses. Therefore, strong adjuvants for stimulating efficient immunologic responses are needed (Harper, 2009).

Currently, marketed vaccines only offer preventive applications, with no therapeutic effects. Thus, designing therapeutic HPV vaccines could be very interesting and beneficial. To devise therapeutic vaccines, post-translational HPV proteins, E6 and E7, could be used as proper candidates. These proteins are consistently present in malignant cervical tumors in high concentrations (Wu et al., 2010). The targeted mechanism in such vaccines is mainly the activation of $CD4^+$ T helper cells (Zajac et al., 1998). However, $CD8^+$ cells have also been proven to play a critical role in vaccine-induced antitumor effect based on the result of studies on animal models (de Oliveira et al., 2015).

On the other hand, a major challenge in the developing and underdeveloped countries is the cost and affordability of the current HPV vaccines, which has a great impact on the high incidence and mortality rates of cervical cancer in such countries (Ferlay et al., 2015). To overcome this issue, the cost-effective production of L2-based vaccines in available prokaryotic microorganisms, such as *Escherichia coli*, has been investigated (Kalnin et al., 2014; Zhai and Tumban, 2016). The second generation HPV prophylactic vaccines, which are L2-based polypeptide vaccines, are proven to be auspicious candidates against the HPV-induced cervical cancer (Tyler et al., 2014). Although L2-based HPV vaccines produce low titers of neutralizing antibodies compared to the first generation L1-based HPV vaccines, they induce wide ranges of cross-neutralizing antibodies against various heterologous HPV types (Schellenbacher et al., 2017).

Multi-epitope polypeptide vaccines are considered as an ideal modality for the prevention and treatment of tumors and cancers (Moyle, 2017), which could be produced in *E. coli*. Although, designing of such vaccines with proper efficacy is quite challenging. The low immunogenicity of L2-based therapeutic vaccines seems to be one of such key challenges (Jiang et al., 2016). To tackle this problem, a variety of vaccine adjuvants have been employed, including alum (Jagu et al., 2011) and Ribi adjuvant (an adjuvant system consisting of monophosphoryl lipid A and trehalose dimycolate) (Jagu et al., 2013; Cargnelutti et al., 2013).Adjuvants that are usually employed in multi-epitope vaccines are originated from pathogen-associated molecular patterns (PAMPs). Toll-like receptors (TLRs) are a universally known family of PAMPs of which 10 types are found in human (Mosaheb et al., 2017). TLR agonists stimulate the adaptive immune responses by inducing the antigen presenting cells (APCs).

A variety of TLR agonists have been applied in previous studies, for instance, alum + MPL (monophpsphoryl lipid A) and alum + CpG (Jagu et al., 2013). In our previous study, the flagellin of *Salmonella enterica*, as an adjuvant with TLR5 agonistic effect, was also used in combination with RS09, as a TLR4 agonist (Negahdaripour et al., 2020). Moreover, some TLR agonists have been employed in the production of approved vaccines (Reed et al., 2016). For instance, the administration of an aluminum salt and Monophosphoryl lipid A (MLA)-containing adjuvant system 04 (AS04), acting as a TLR4 agonist, in Cervarix® structure (Jiang et al., 2016), was proven to produce a more potent immunologic effect than the aluminum salt alone (Giannini et al., 2006).

The use of reverse vaccinology approaches *via* bioinformatics tools plays a major role in novel vaccine development, considering the fact that the use of such tools has become increasingly popular in recent years.

In this study, immunoinformatics tools are utilized to facilitate the epitope selection for designing two different HPV candidate vaccine constructs with both prophylactic and therapeutic properties, which can promote both humoral and cellular immune responses. Evaluation of various properties of these constructs are done by various computational tools to identify the best construct with proper physiochemical properties and immunogenicity.

2. Methods

The methodology of this project could be summarized in several steps:

1. Identification, selection, and evaluation of different antigenic epitopes.
2. Designing a linear combination of selected epitope sequences and joining them *via* linkers.
3. Evaluation different physiochemical characteristics of the designed multi-epitope vaccine.

2.1. Sequence retrieval

The universal protein resource (UniProt) database (https://www.uniprot.org/) was used to retrieve the protein sequences of HPV16 L2 (UniProt ID: P03107), E6 (UniProt ID: P03126), E7 (UniProt ID: P03129), *Mycobacterium tuberculosis* RpfB (UniProt ID: P9WG29), *M. tuberculosis* RpfE (UniProt ID: O53177) and human TLR-4 (UniProt ID: O00206) (UniProt, 2021).

2.2. Immunoinformatics analysis

Novel approaches in vaccinology administer various immunoinformatic tools and databases to supply well-established data input, which are employed to achieve high throughput and methodologically supported candidate vaccine designed constructs (Hegde et al., 2018) (Fig. 1).

2.2.1. MHC-I epitope prediction

The MHC-I epitope of L2, E6, and E7 proteins were identified using two bioinformatics servers: nHLApred (http://crdd.osdd.net/raghava/nhlapred) and IEDB (http://tools.iedb.org/mhci). Different human MHC-I alleles were selected based on their population distribution (HLA-A1, HLA-A2, HLA-A1101, HLA-A24, HLA-A3, HLA-B0702, HLA-B8), and the length of epitope sequences was set at 9-mer.

nHLApred uses an algorithm of epitope prediction based on a combination of artificial neural networks (ANNs) and quantitative matrix-based methods (Lata et al., 2007); whereas IEDB server employs IEDB recommended method in which a consensus method consisting of ANN (Lundegaard et al., 2008; Nielsen et al., 2003) stabilized matrix method (SMM) (Peters and Sette, 2005), and scoring matrices derived from the combinatorial peptide library (Comblib sidney2008) (Sidney et al., 2008) or NetMHCpan is used (Kim et al., 2012).

2.2.2. Prediction of CTL epitopes and peptide-MHC interfaces

Prediction of CTL epitopes was carried out by CTLPred server (http://crdd.osdd.net/raghava/ctlpred/), which uses a direct method of prediction for CTL epitopes. CTLPred utilizes ANN and support vector machine (SVM) methods separately or a combination of them (Soria-Guerra et al., 2015). CTLPred server considers a cut-off score to differentiate epitope sequences from non-epitope ones. In this study, the default cut-off score and the combined method were used.

2.2.3. MHC-II binding epitope prediction

Two bioinformatics servers, ProPred and MHC2Pred, were utilized to identify the MHC-II binding epitopes of L2, E6, and E7 proteins. Different human MHC-II alleles were selected based on the distribution and population of the alleles (DRB1–0101, DRB1–0301, DRB1–0401, DRB1–0405, DRB1–0701, DRB1–1101, DRB1–1301, DRB1–1501, DRB5–0101). ProPred (http://crdd.osdd.net/raghava/propred) employs the quantitative matrix method for MHC-II epitope prediction,

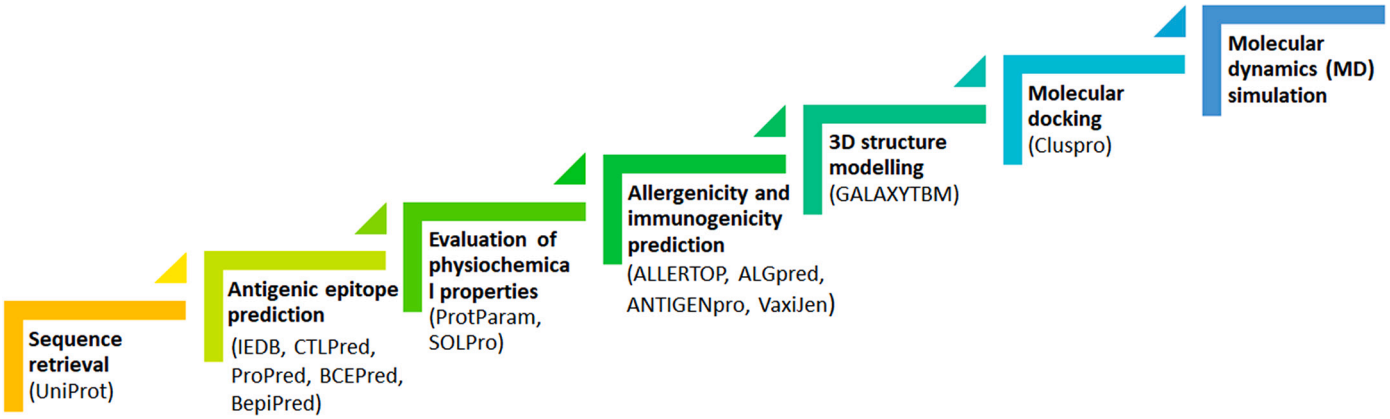

Fig. 1. An overview of performed vaccine design process using immunoinformatic tools.

whereas MHC2pred (http://crdd.osdd.net/raghava/mhc2pred) uses the SVM-based methods, which enables to improve the selections (Lata et al., 2007).

2.2.4. Linear B cell epitope prediction

The selection of B cell epitopes was carried out by two bioinformatics servers: BepiPred (http://www.cbs.dtu.dk/services/BepiPred) and BCEPred (http://crdd.osdd.net/raghava/bcepred), both of which utilize a prediction method of continuous B cell epitopes based on amino acid properties such as hydrophilicity, charge, and exposed surface area (Soria-Guerra et al., 2015). BepiPred also uses a combination of a hidden markov model and propensity scale method (Larsen et al., 2006). The accuracy of the B-cell prediction method employed by BCEPred is 58.70%, and this server combines four residue properties to enhance prediction accuracy (Saha and GPS, 2004).

2.2.5. Selection of epitope sequences

Selection of epitope sequences was achieved by choosing different types of epitopes discussed before (MHC-I, MHC-II, CTL, and linear B-cell epitopes). After assessing different epitope sequences showing the most amino acid overlaps, a number of the epitopes that could stimulate the immune responses with the highest probability were selected.

2.3. Evaluation of the selected epitopes and designing a vaccine construct

2.3.1. Vaccine construction and prediction of physiochemical parameters of vaccine

Based on the obtained results, two different vaccine constructs were designed. One of them contained the same TLR agonist at both terminals, and the second one had two different TLR agonists at the two terminals of the peptide construct. For the first construct design, two G5 domains of RpfB protein from *M. tuberculosis* were used at the C-terminal and N-terminal of the construct. On the other hand, for the second construct, RpfE protein (from *M. tuberculosis)* was administered at the N-terminal of construct, and a G5 domain of RpfB protein was attached to the C-terminal. Two different arrangements of the linear multi-epitopes were designed as candidates in order to reach the best possible immunogenicity and physiochemical factors after the assessment of the aforesaid factors.

Solubility of the designed constructs was evaluated using the SOLpro (http://scratch.proteomics.ics.uci.edu/), which predicts the solubility of a given protein after overexpression in *E. coli* utilizing a two-step SVM architecture based on multiple different representations of the primary peptide sequence. SOLpro server has a 74% prediction accuracy achieved by multiple runs of a 10-fold validation (Magnan et al., 2009).

In order to predict other physiochemical properties of the designed linear construct, the ProtParam online tool (https://web.expasy.org/protparam/) was utilized. ProtParam computes different physiochemical properties that could be inferred from different protein sequences, including the molecular weight (MW), isoelectric pH (pI), extinction coefficient, instability index, aliphatic index, and grand average of hydropathicity (GRAVY) (Gasteiger et al., 2005).

2.3.2. Allergenicity prediction

To predict the allergenicity, two prediction tools were utilized to minimize false prediction. AllerTop v. 2.0 (https://www.ddg-pharmfac.net/AllerTOP/) is a server used for the *in silico* prediction of allergens. It computes data using a method based on auto cross covariance (ACC). The performing method also utilizes a machine learning classification model called k-nearest neighbors (kNN), with an accuracy of 85.3% at 5-fold cross-validation (Dimitrov et al., 2014). The web server AlgPred (http://crdd.osdd.net/raghava/algpred/) was developed for the prediction of allergenic proteins and IgE-epitope mapping on proteins. Computation of allergen prediction by AlgPred is carried out by a standard SVM-based method, presence of IgE epitopes, motif-based prediction (MEME/MAST), prediction of the ARPs, and a hybrid method that utilizes all the former methods to achieve better and reliable results (Saha and GPS, 2004).

2.3.3. Antigenicity prediction

Antigenicity is one of the main properties of proteins that should be evaluated in vaccine design. ANTIGENpro and Vaxijen v2.0 were applied in this study to achieve better prediction accuracy. Vaxijen v2.0 (http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html) is an online prediction server, which uses a method based on auto-cross covariance transformation of protein sequences into uniform vectors of principal amino acid properties to evade the limitations of alignment-dependent methods (Doytchinova and Flower, 2007). In this research, virus was selected as the target organism, and the threshold was set at 0.4 ANTIGENpro (http://scratch.proteomics.ics.uci.edu/) is an alignment-free and pathogen-independent predictor for prediction of protein antigenicity. ANTIGENpro uses a two-step prediction model consisting of multiple representations of the primary sequence and five machine learning algorithms resulting in a high-throughput antigenicity prediction (Magnan et al., 2010).

2.4. Tertiary structure

2.4.1. Homology modeling

Homology modeling was used for building the 3D structure of linear amino acid sequences of the designed vaccine construct. 3D modeling was carried out by GalaxyWEB server (http://galaxy.seoklab.org/cgi-bin/submit.cgi?type=TBM), which performs a high-accuracy prediction of the tertiary protein structure. Its algorithm consists of four steps: identification of similar protein templates, iterative structure assembly, function annotation structure selection, and refinement of the model (Ko

et al., 2012a).

The method utilized by GalaxyTBM server has been tested in recent community-wide CASP9 experiments (9th Critical Assessment of techniques for protein Structure Prediction), and the server has been ranked as one of the best servers for protein structure prediction (Ko et al., 2012b). The selected output model of GalaxyTBM was visualized by including Drug Discovery Studio 3.5 and PyMol visualizer.

The 3D model of TLR4 was obtained from the SWISS-model data bank (https://swissmodel.expasy.org/). TLR4 is a homodimer receptor (Krüger et al., 2017). In order to dock the designed model with TLR4 receptor, the surrounding water molecules and one of the homodimer monomers were eliminated.

2.4.2. Validation of the 3D models

The obtained 3D models were validated using ProSA-web (Protein structure analysis), Procheck, and ERRAT web servers. ProSA-web (https://prosa.services.came.sbg.ac.at/prosa.php) is an online tool that was used for validating the 3D protein model candidates. ProSA-web compares the quality of the experimental 3D protein models with all the known protein structures available in the PDB database using quality *Z*-score based on the X-Ray analysis and NMR spectroscopy (Wiederstein and Sippl, 2007). ERRAT server (https://servicesn.mbi.ucla.edu/ERRAT/) measures and analyses non-bonded interactions and evaluates the errors in the spatial conformation of the protein model (Colovos and Yeates TO, 1993). Procheck is another online server that evaluates the stereochemical quality of the proteins by presenting Ramachandran plots of the amino acid residues, which could further evaluate the spatial conformation of the protein model. The Ramachandran plot evaluates the phi-psi torsion angles of amino acid residues and categorizes the residues in the most favored, additionally allowed, generously allowed, and disallowed regions (Laskowski et al., 1993).

2.4.3. Prediction of conformational B-cell epitopes

Discotope 2.0 is an online bioinformatics tool (http://www.cbs.dtu.dk/services/DiscoTope/) predicting the discontinuous linear B-cell epitopes that are present in the 3D protein model. Discotope 2.0 utilizes a combined score method using proximity summed log-odds ratios and surface measures of the epitopes to compute an overall score (Kringelum et al., 2012). In this study, the default threshold setting of −3.7 was employed.

2.5. Prediction of TLR4/MD2 interaction with vaccine candidate

2.5.1. Protein-protein docking and the evaluation of docked models

In order to predict the protein-protein interactions between the vaccine candidate and TLR4, ClusPro server (https://cluspro.org) was utilized. ClusPro uses three steps to compute the protein-protein interactions between the two proteins: 1) rigid body docking, 2) root-mean-square deviation (RMSD) based sorting of the top 1000 models with the lowest energy index, and 3) energy minimization of the selected model (Kozakov et al., 2017). In order to achieve the best protein-protein interaction model, different physiochemical factors such as hydrophobicity, electrical charge distribution, and amino acid residues in the interaction site should be taken into consideration.

Following the determination of protein-protein interactions between the vaccine construct and the TLR4/MD2 complex by Cluspro, the output models of Cluspro were further evaluated by PyMol viewer in order to assess if the output models possess the desired interactions or not.

2.5.2. Molecular dynamics simulation

Although molecular docking assesses the interactions involved in protein-ligand binding, changes regarding these interactions should be evaluated in a time frame, as protein-ligand interactions similar to any other biological and molecular event, is a dynamic matter rather than a static one. The behavior and stability of biological macromolecules during a time frame can be assayed by molecular dynamics (MD) simulation, which is a powerful technique in bioinformatics approaches (Boehr et al., 2009; Hansson et al., 2002).

In this study, MD simulation tool was used to monitor the stability of the designed vaccine binding to TLR4/MD2 complex during 90 ns.

The initial configuration of MD simulation was obtained by the docking process. In fact, the chosen docked model of TLR4/MD2-vaccine complex was solvated in a triclinic box of TIP3P water molecules. GROMACS 5.0.1 package was used to perform MD simulation process, while all conditions similar to our previous studies were applied (Vakili et al., 2018; Negahdaripour et al., 2017).

3. Results

3.1. Sequence retrieval and assessment

After acquiring amino acid sequences for L2 (P03107), E6 (P03126) and E7 (P03129), RpfB (P9WG29) and RpfE (O53177) from UniProt server, the physiochemical properties of the sequences were also evaluated using Protparam server. Notably, just 120 amino acids in the N-terminal of L2 protein were employed according to the exposed domain of the protein. Moreover, only the G5 domain of RpfB protein was used in this study (Table S1).

3.2. Immunoinformatics evaluation

3.2.1. MHC-I binding epitope prediction

nHLApred and IEDB servers presented 9-mer MHC-I binding epitopes from the N-terminal end of L2, E6, and E7 proteins for seven human MHC-I alleles (HLA-A1, HLA-A2, HLA-A1101, HLA-A24, HLA-A3, HLA-B0702, HLA-B8). The overlapping sequences of the predicted epitope sequences for each allele were selected to achieve epitope sequences with the highest probability of immunogenicity (Table S2).

3.2.2. CTL epitope and linear B cell epitope prediction

CTL epitope prediction was carried out by CTLpred and PAComplex servers, which resulted in the identification of CTL binding epitope in L2, E6, and E7 proteins. These epitopes were ultimately limited by selection of the overlapped predicted regions between the two aforesaid servers. Linear B cell epitopes were also predicted by BCEpred and BepiPred 2.0 servers (Table S2).

3.2.3. Prediction of MHC-II binding epitopes

ProPred and MHC2Pred servers were utilized in order to identify the MHC-II binding epitope regions present in L2, E6, and E7 protein sequences for nine MHC-II human alleles (DRB1–0101, DRB1–0301, DRB1–0401, DRB1–0405, DRB1–0701, DRB1–1101, DRB1–1301, DRB1–1501, DRB5–0101). These predicted regions helped us for selecting the overlapping regions between different MHC-II alleles (Table S2).

3.2.4. Selection of the final epitope sequences

The overlapped regions were selected based on the results achieved in the above analyses. The results were compared, and ultimately four epitope segments, two from L2 protein and one from each of E6 and E7 proteins, were selected (Table 1).

3.3. Designing vaccine construct and evaluation of the selected sequences

3.3.1. Vaccine construct design

The selected epitope segments from each protein, which were identified in the previous section, 12–23 and 78–89 amino acid segments from L2 protein, 93–106 amino acid segment from E6 protein, and 70–94 amino acid segment from E7 protein, were employed in the vaccine construct based on the analysis results and previous studies.

Table 1
The selected epitope sequences used for designing two different vaccine constructs.

Protein	Start	End	Sequence	Algpred	AllerTop 2.0	Vaxijen	Antigenpro
L2	12	23	RASATQLYKTCK	Non-allergen	Non-allergen	Antigen	Antigen
L2	78	89	TRPPTATDTLAPV	Non-allergen	Non-allergen	Antigen	Antigen
E6	11	27	DPQERPRKLPQLCTELQ	Non-allergen	Non-allergen	Antigen	Antigen
E7	70	89	QSTHVDIRTLEDLLMGTLGIVCPIC	Non-allergen	Non-allergen	Antigen	Antigen

These sequences were inserted in the construct along with the G5 domain of RpfB protein and RpfE protein as adjuvants, which are TLR4 agonists (Kim et al., 2013; Faridgohar and Nikoueinejad, 2017; Kumar et al., 2019). TLR4 agonist adjuvants were utilized in the N-terminal and C-terminal of the vaccine construct resulting in two different linear constructs, a 294- amino acid linear construct consisting of RpfB protein at the N- and C-terminals of the construct (construct A), and a 357-amino acid linear construct with RpfE protein at N-terminal and RpfB at C-terminal of the linear construct (construct B). The epitope segments and the adjuvants were linked using $(EAAAK)_3$ and GGGGS linkers. Both designed constructs are shown in Fig. 2.

3.3.2. *Physicochemical, allergenicity, and antigenicity evaluation of the constructs*

The assessment of physicochemical features (pI, aliphatic index, instability index, MW and GRAVY) of the linear protein constructs was carried out by the ProtParam server. Solubility of the constructs was also predicted by the Solpro online tool. The results are available in Table 2.

Allergenicity of the whole peptide constructs were predicted using Allertop v.2.0 and Algpred servers, as shown in Table 2. Moreover, antigenicity of the designed constructs was predicted by AntigenPro and VaxiJen v2.0 servers (Table 2).

3.4. *Tertiary structure analysis*

3.4.1. *Homology 3D modeling and validation of models*

The 3D models of the designed protein candidates were predicted by GalaxyTBM server based on homology modeling. GalaxyTBM presents the five top 3D models based on HHsearch score, which indicates the model quality after model refinement carried out by the GalaxyTBM during modeling process. After homology modeling, the designed models were validated using ERRAT, ProSA-Web, and Ramachandran plot in order to select the best predicted model between GalaxyTBM output models. The 3D vaccine models are presented in Fig. 3a and b for

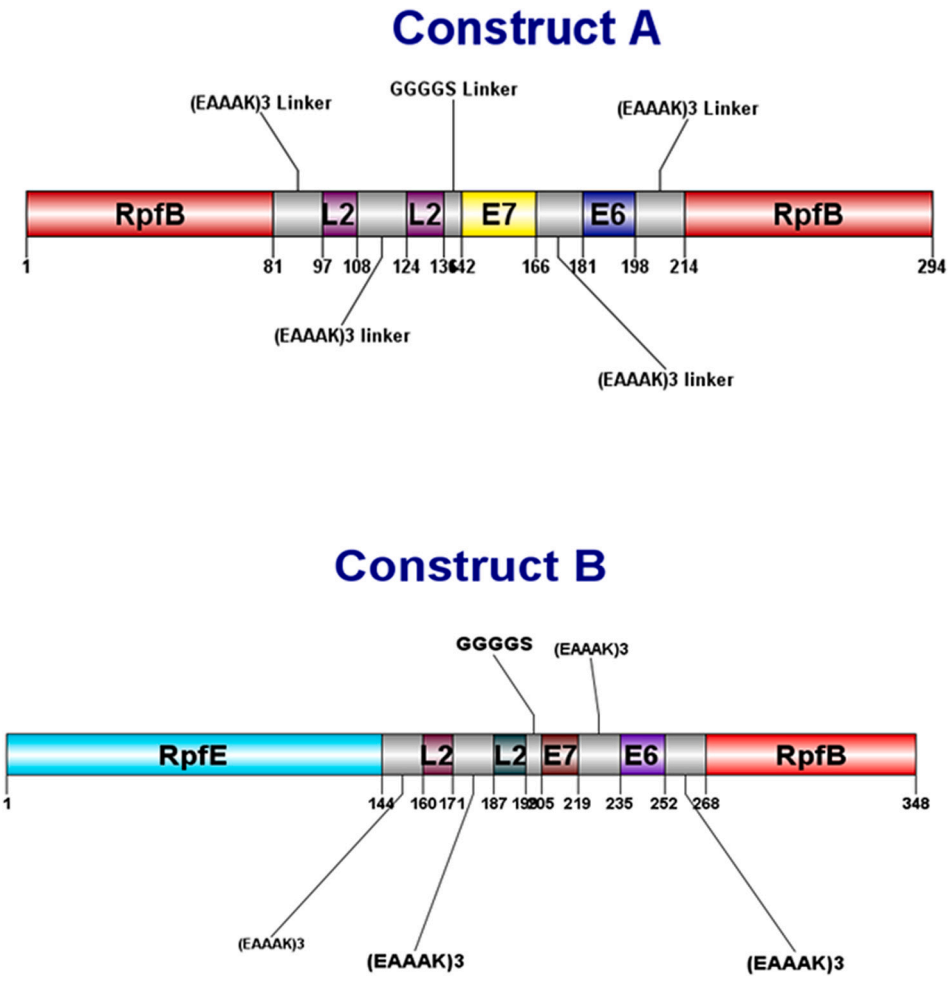

Fig. 2. Schematic design of the vaccine constructs a) Construct A: Consisting of two RpfB G5 domain as adjuvant, two L2-based epitope regions, an E6-based epitope, and an E7-based epitope region, linked together by various peptide linkers. b) Construct B: This construct is similar to construct A, but contains another adjuvant, RpfE, which is inserted at the N0-terminal of the construct (instead of RpfB G5 domain at N-terminal of construct A0.

Table 2
Physicochemical properties of the designed vaccine constructs.

	Vaccine construct A (2RpfB)	Vaccine construct B (RpfB + RpfE)
Physicochemical properties		
Solubility (SOLpro)	0.741185	0.683546
Molecular weight	30,948.35	36,929.38
Theoritical pI	7.79	5.08
Instability index	39.61	43.58
Extinction coefficient	1740	28,335
Aliphatic index	80.71	71.57
GRAVY	−0.353	−0.397
Number of Amino acids	294	357
Total number of negatively charged residues (Asp + Glu)	41	48
Total number of positively charged residues (Arg + Lys)	42	39
Allergenicity		
AllerTOP v. 2.0	Non-allergen	Non-allergen
Algpred	Probable non-allergen	Probable non-allergen
Antigenicity		
ANTIGENpro	Probable antigen 0.823322	Probable antigen 0.909710
Vaxijen v. 2.0	Probable antigen 0.4735	Probable antigen 0.4039
B cell epitope/Discotope 2.0	88/294	116/357

construct A and B, respectively. The results of the 3D model validation servers for both A and B constructs are shown in Figs. S1 and S2, respectively.

3.4.2. Prediction of B-cell conformational epitopes

Discotope 2.0 server was employed to predict conformational B-cell epitopes. For construct A, 88 conformational epitopes out of 294 residues were discovered, while for construct B, 116 out of 357 residues were identified (Table 2).

3.5. Protein-protein interaction exploration

3.5.1. Docking of vaccine construct with TLR4 protein

In order to assess the interactions between the vaccine construct candidate and TLR4, molecular docking of the 3D models of the mentioned proteins was carried out by the Cluspro server. Due to the domination of hydrophobic interactions in TLR4/MD2-TLR4 agonist binding and the hydrophobic pocketing of the TLR4 agonist inside of MD2 protein of the TLR4/MD2 receptor complex as shown in previous studies (Park et al., 2009), the results were filtered by the hydrophobic dominant interactions by applying the hydrophobic-favored filter on server result page to achieve more realistic results. The best docking model was selected based on the above- mentioned factor (Fig. 4).

3.5.2. Molecular dynamic simulation (MD)

Firstly, the equilibration and production MD steps were assayed to ensure the accuracy of the MD simulation steps. In this regard, some parameters such as pressure, temperature, volume, density, and energy values were evaluated. All parameters confirmed that MD simulation process was performed correctly. In the following, the MD generated trajectory was used to extract some important parameters. These parameters were assayed to monitor the behavior of TLR4/MD2-vaccine complex during the MD simulation time. TLR4, MD2, and vaccine

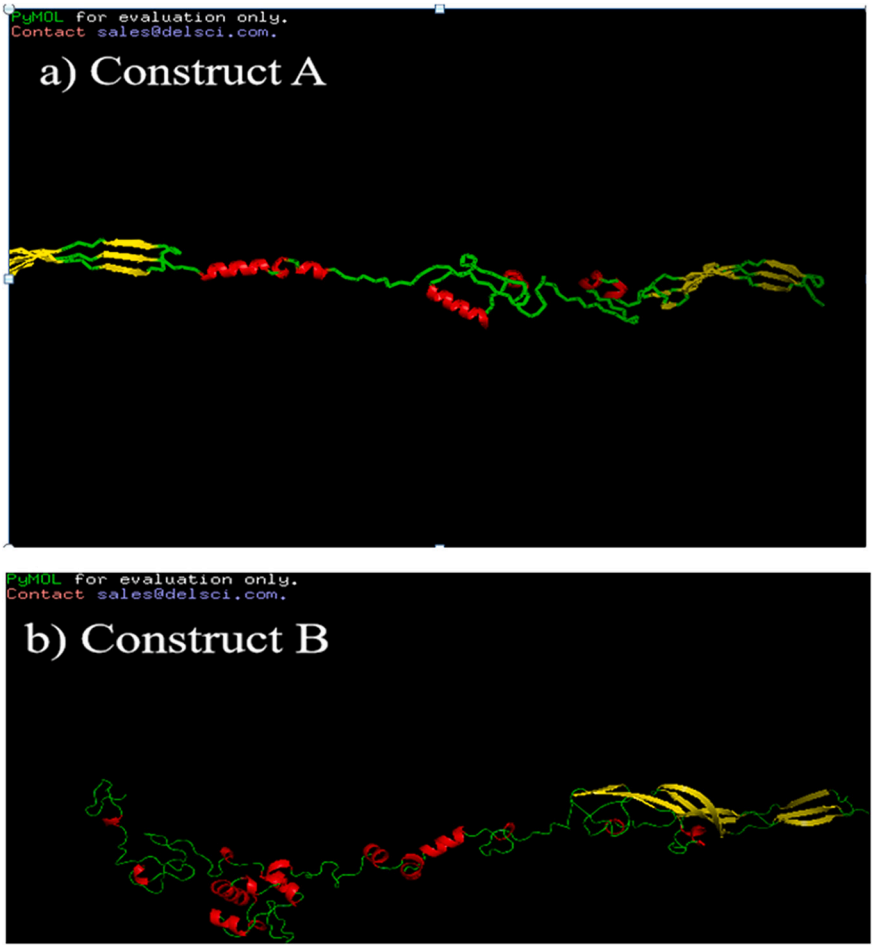

Fig. 3. The 3D vaccine structures obtained by GalaxyTBM server. a) Construct A, b) Construct B.

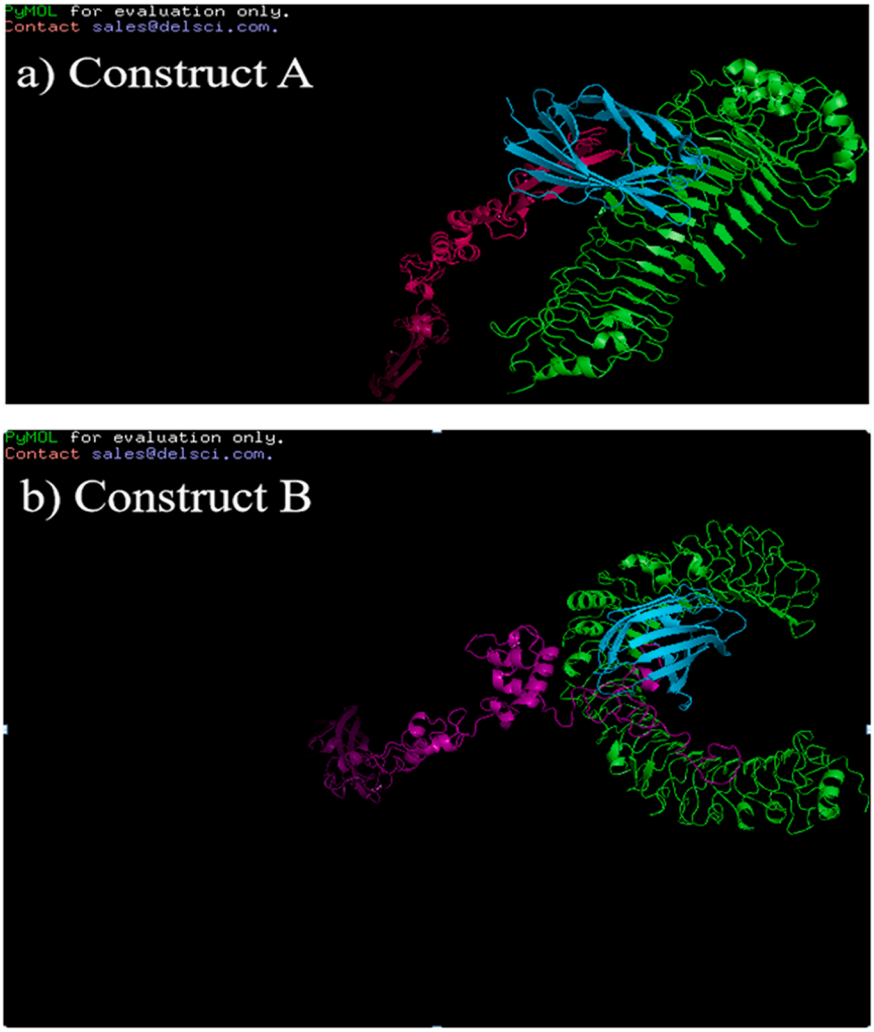

Fig. 4. Molecular docking results. a) Protein construct A, b) Protein construct B.

molecules are given proper opportunity during 90 ns of MD simulation to obtain their best interactions and achieve the most appropriate orientation relative to each other.

In order to evaluate the stability of the backbone atoms, root mean square deviation (RMSD) parameter was evaluated for TLR4, MD2, and the vaccine protein chains. The principal receptor of the vaccine molecule is MD2 protein, although the final structure obtained at the end of the simulation time showed there were a few intermolecular interactions between TLR4 and vaccine molecules. TLR4 and MD2 proteins interacted with each other *via* some strong intermolecular interactions as well. RMSD plots of TLR4 and MD2 proteins show that they had appropriate stability resulting from their intermolecular and intramolecular interactions (Fig. 5). It should be noted that these proteins underwent insignificant fluctuations during MD simulation time due to their efforts to obtain the best position relative to each other and the vaccine molecule. As shown in Fig. 5, the vaccine molecule suffered significant fluctuations, especially from the beginning of the simulation to the 40th ns of the simulation time, although it almost reached a steady

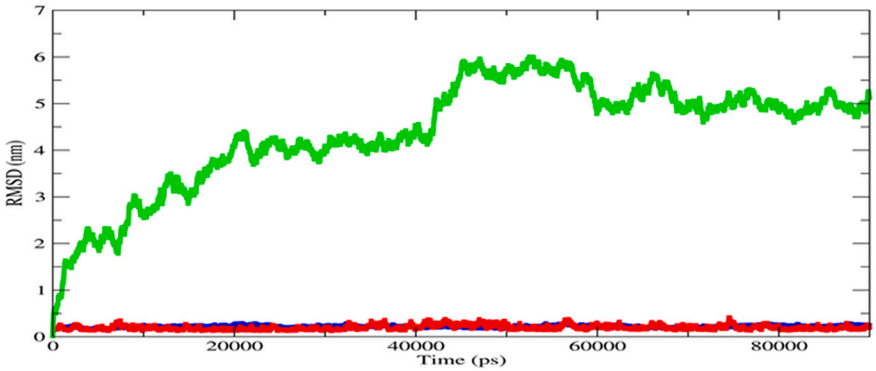

Fig. 5. RMSD plot during molecular dynamics simulation for the vaccine construct (A) (green), TLR4 (red), and MD2 (blue) molecules. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

state from the 60th ns on. It seems that this molecule endured some displacements to obtain a stable position relative to the two other proteins. In fact, MD simulation provided a special opportunity for TLR4/MD2-vaccine complex to move to the most stable situation.

In the next step, changes in radius of gyration (R_g) were assayed to evaluate the compactness of the protein chains over the MD simulation time. Rg plots of TLR4, MD2, and vaccine molecules are shown in Fig. 6. Changes in the Rg plots of all protein chains are almost identical to the RMSD plots. This matching is the golden point of the MD simulation process. As mentioned above, the backbone atoms of the vaccine molecule suffered significant displacements until 40th ns; and they reached a steady state since the 60th ns of the simulation time. The same trend was seen in the Rg plots. The Rg plot of the vaccine molecule clarifies that the observed displacements in the RMSD plot were arisen from the tendency of the vaccine molecule to get compact during the MD simulation time. Strong intramolecular interactions have almost kept the Rg values of MD2 and TLR4 proteins, although the shape and intermolecular interactions of the MD2 protein with the vaccine molecule led to decrease the fluctuations of its RMSD and Rg parameters in comparison with TLR4 protein as well.

In the following, changes of the hydrogen bond number over the MD simulation time were also evaluated. A complete concordance was seen between the changes of hydrogen bond number and the changes of the two previously-mentioned parameters (Rg and RMSD parameters). According to Fig. S3, number of hydrogen bonds between the vaccine molecule and MD2 protein suffered some fluctuations during the MD simulation time. In particular, this parameter had the lowest values from the 40th ns to the 50th ns of the simulation time. The RMSD plot reached to the highest values over this period; on the contrary, the Rg plot got the lowest values during the aforesaid 10 ns. All these results confirm that the three molecules tried to get the best position relative to each other over the 90 ns simulation time.

The root-mean-square fluctuations (RMSF) parameter of all residues of TLR4, MD2, and vaccine molecules were computed as well. The majority of TLR4 and MD2 residues underwent small fluctuations during the MD simulation time resulting from powerful intermolecular and intramolecular interactions (Fig. 7). The N/C terminal residues of TLR4 protein and a few residues of this protein were located far from MD2, and the vaccine molecule suffered from high fluctuations (Fig. 7.a). The N158 residue of the MD2 protein and a few residues of this protein that were placed far from TLR4 and vaccine molecules also endured high fluctuations (Fig. 7.b). These residues were exposed to the solvent, so they experienced more movements than other residues over the MD simulation time.

The three-dimensional representation of the interaction interface between MD2 protein and the designed vaccine is shown in Fig. 8. F126, I54, Y131, I124, R90, K91, S118, K122, and some other residues were already mentioned as the main residues of MD2 interacting with an agonist (Park et al., 2009; Carpenter and O'Neill, 2009; Billod et al., 2016). The designed vaccine molecule engaged with the majority of the mentioned residues appropriately. It was previously shown that the F126 residue of the MD2 protein and its conformational change play a critical role for the binding of this protein to its agonist. This residue adopted a closed state situation, when the protein interacted with the cognate agonist (Carpenter and O'Neill, 2009; Billod et al., 2016). The conformation of this residue was specified in the crystallographic structure as well as the obtained structure at the end of simulation time (Fig. 9). The closed state of this residue was maintained in the presence of our vaccine molecule over the 90 ns of MD simulation time.

4. Discussion

In the recent years, advances in bioinformatics tools, genomics, and biological databases have made a significant impact on designing vaccines against various infectious diseases and cancers (María et al., 2017). Due to considerable mortality and new-case infections of HPV, preventive and therapeutic strategies are of utmost importance. Development of a vaccine against HPV with dual therapeutic and prophylactic properties could be an effective way to control cervical cancer. Despite many studies on prophylactic or therapeutic HPV vaccines, the dual approach is less found in previous research. Zhao et al. recently reported administration of a nanoparticle dual purpose HPV vaccine in mice, which protected the mice against tumor development upon two subsequent tumor challenges. Their developed vaccine was a novel design containing L2- and E7-specific epitopes inserted on thioredoxin (Trx) scaffold surface. A heptamerization-promoting module (OVX313) was also employed for production of nanoparticles (Zhao et al., 2020). TA-CIN is also a recombinant HPV16 L2/E6/E7 fusion protein vaccine, which showed optimum results in preclinical studies (Van der Burg et al., 2001) and is presently in an ongoing Phase 1 clinical study (NCT02405221).

To achieve prophylaxis, L2-based vaccines have been studied in multiple studies with different adjuvants or antigens. In a study by Schellenbacher et al., utilization of RG1 epitope, a region harvested from L-2 protein (aa17–36), along with HPV16 VLPs showed acceptable results in regards to achieving a broad-spectrum immunogenicity against both cutaneous and mucosal HPVs. These results were also assessed using Pseudoviron-based neutralization assay and harvesting neutralizing antibodies against various types of HPVs in mice an rabbits (Schellenbacher et al., 2013). Moreover, intranasal administration of cross-neutralizing epitopes from L2 capsid protein without immunogenic adjuvants in mice demonstrated promising results but failed to generate proper immunogenicity in a phase 1 clinical trial (Kawana et al., 2001). In our previous study, flagellin and a short synthetic TLR4

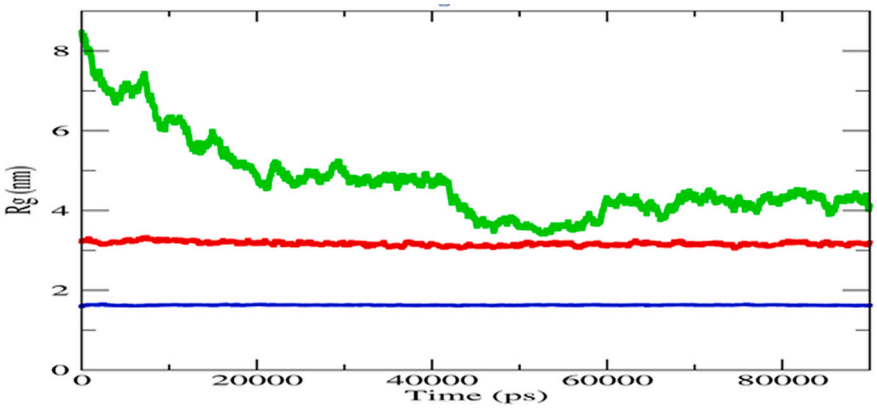

Fig. 6. Rg plot during MD simulation for the vaccine construct (A) (green), TLR4 (red), and MD2 (blue) molecules. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

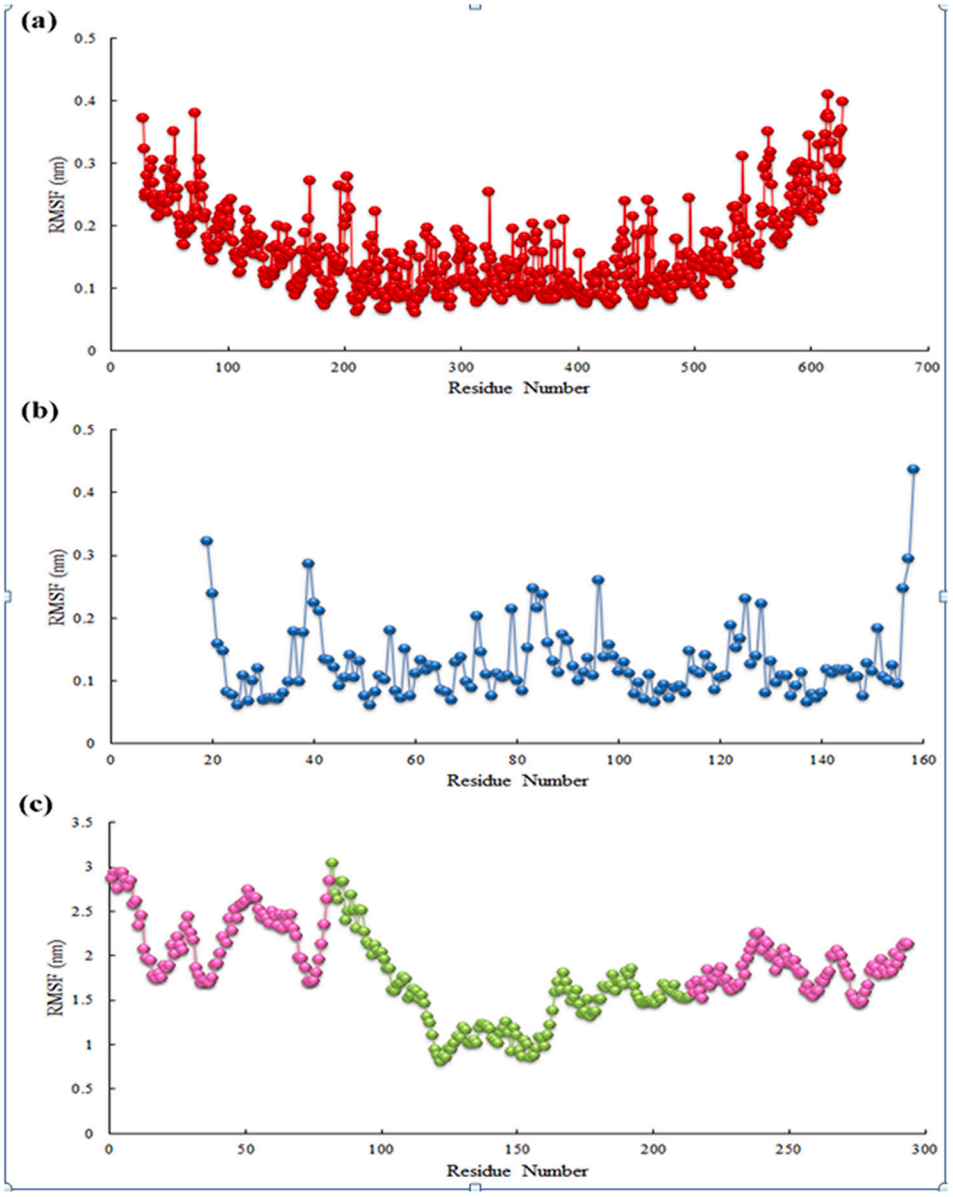

Fig. 7. RMSF plot of all residues of: (a) TLR4 protein, (b) MD2 protein, and (c) vaccine molecule (residues of RpfB adjuvants are colored in magenta). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

agonist was employed along with L2 epitopes as a prophylactic vaccine. The vaccine administration in mice led to cellular and humoral immune responses, with a more Th-1 favored trend (Negahdaripour et al., 2020).

L2-based prophylactic HPV vaccines are yet to pass clinical trials and enter the market due to the low immunogenicity displayed by the L2 capsid protein (Schiller and Lowy, 2015).

E6 and E7 proteins, which are early post-translation proteins, are also tailored targets for therapeutic vaccines due to their function in the viral pathogenicity (Zur, 2002). A variety of therapeutic vaccines using E6 and E7 proteins have been designed and entered the clinical trials. A phase I clinical study on an E6- and E7-based vaccine showed appropriate immunogenicity with low toxicity in the end-stage cervical cancer patients (Kenter et al., 2008). In a similar study, the results of a therapeutic E6- and E7-based vaccine on papillomavirus rabbit models, were proved to be promising (Vambutas et al., 2005). The efficacy of an E6- and E7-based-vaccine was confirmed to correspond with the T-cell immune responses induced by the vaccine, as shown in a clinical study (van Poelgeest et al., 2016). However, a clinical trial regarding the efficacy, safety and tolerability of an E6/E7 therapeutic vaccine coinjected with IL-12 plasmid has been carried out by *Inovio Pharmaceuticals* from 2014 to 2017 (NCT02172911. Phase 1 results of this vaccine candidate showed cleared biopsies in patients after chemoradiation and vaccination and also indicated the safety of this therapeutic candidate (Hasan et al., 2020). Although the safety and immunogenicity of these vaccines have been confirmed, no therapeutic vaccine candidate is approved for clinical use yet.

In this study, resuscitation-promoting factor E (RpfE) and the G5 domain of RpfB from *M. tuberculosis* were employed as TLR4 agonists to design two different vaccine construct candidates. RpfB induces the Th-1 cell type cellular immunity by TLR-dependent activation of dendritic cells, which subsequently promotes the expression of cell surface molecules, such as CD80, CD86, MHC-I, and MHC-II, along with pro-inflammatory cytokines (TNF-α, IL-12p70, IL-6, and IL-β). Moreover, it was shown that the G5 domain of RpfB plays an essential role in the interaction of RpfB and TLR4 (Kim et al., 2013). Moreover, RpfE protein has shown prominent stimulatory effects on dendritic cell maturation by activating mitogen-activating protein kinase (MAPK) and nuclear factor kappa B (NF-κB) thus acting as a potential vaccine adjuvant (Choi et al.,

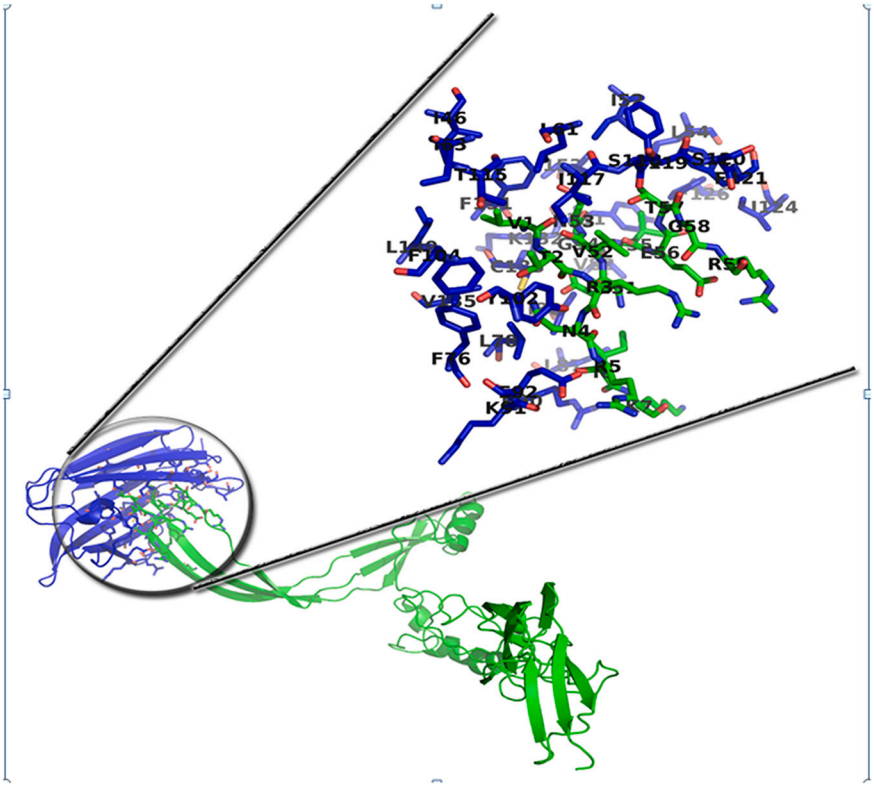

Fig. 8. Three-dimensional representations of the interaction interface between the MD2 (blue residues) and vaccine (green residues) proteins at the end of the MD simulation time. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

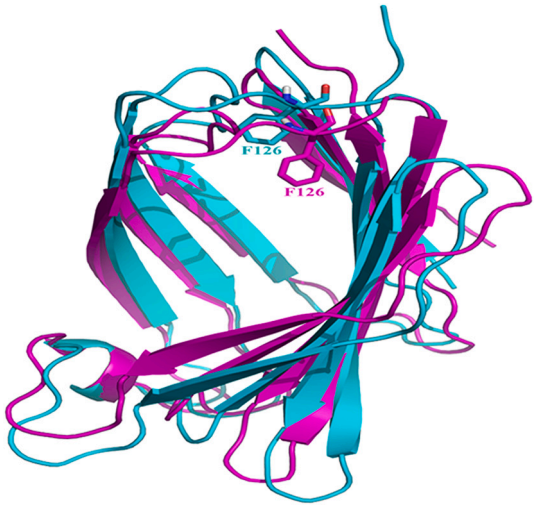

Fig. 9. Superimposition of the MD2 protein chain at the end of simulation time (cyan chain) over its structure obtained at the docking process (magenta chain). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2015). RpfE also incites the differentiation of naïve T-cells to Th-1 and Th-17 cells, which then contributes to inducing the expression of cell surface molecules and pro-inflammatory cytokines (IL-12p70, IL-23p19, IL-6 and Il-β) (Choi et al., 2015).

The TLR4 agonist sequences were employed as built-in adjuvant in our two designed constructs, since the insertion of adjuvants inside the peptide construct, was shown to improve the immunogenicity of the whole construct in comparison to their separate usage (Huleatt et al., 2007).

To achieve the best immunogenicity results in designing both A and B constructs, different epitope identification servers were utilized to find the overlapped sequences. The selected sequences are immunodominant domains, which are able to induce cellular as well as humoral immune responses with the highest possibility. In order to attain the best immunologic epitope sequences for each protein, different servers that use varied predication algorithms were employed to increase the prediction accurateness. Ultimately, the final epitope sequence for each of L2, E6, and E7 proteins was achieved by identifying the overlapping sections of MHC-I, MHC-II, B-cell, and CTL epitope sequences (Table 1).

Given the explosibility of the N-terminus 120 amino acid portion of L2 on the viral surface (Doorbar et al., 2015), candidate epitopes were selected only from this region of the L2 protein. Finally, amino acids 12–23 and 78–89 from L2 protein N-terminus, amino acids 70–89 from E7 protein N-terminus, and amino acids 11–27 from E6 protein N-terminus were chosen as the final epitope segments (Table 1).

Following the selection of epitope sequences, two peptide linkers with different properties were applied to adhere the epitope sequences to the N-terminal and C-terminal adjuvants. Different linkers with variable lengths, possess different properties and can impact protein stability, flexibility, domain-domain orientation, and protein conformation (Doorbar et al., 2015; van Leeuwen et al., 1997). These non-immunogenic motifs facilitate antigen presenting processes and avoid changes in the structure of antigens due to the formation of neoepitopes (Livingston et al., 2002). Flexible linkers, such as GGGGS, contain two small polar and non-polar amino acids, glycine (G) and serine (S), which provide solubility and flexibility in the protein structure due to their small sizes (Chen et al., 2013; Argos, 1990). On the other hand, EAAAK peptide linker, provide stability in the protein sequence by applying its rigid nature in joining different segments (Arai et al., 2004). In our construct, the use of $(EAAAK)_3$ linker helped to balance the acidic pI of some of the selected epitope regions and bring the pI of the construct near to the physiological pH levels. In multi-epitope vaccine designing, the position of selected epitopes also affects the physiochemical

properties of the peptide and the 3D folding of the protein, which contributes to the vaccine efficacy (Negahdaripour et al., 2018). It should be noted that many different constructs could be built using these epitopes with diverse orders. Given that evaluation all these possibilities one by one is not feasible, these tow constructs were selected among several tested ones (data not shown) based on their better physiochemical and immunological properties as exhibited by the aforementioned programs and servers. Development of computational tools for evaluation of various possible structures would be a great help to overcome this limitation of vaccine designing.

Following construction of the sequences, the physiochemical properties of the constructs were evaluated and compared by Protparam and SolPro (Table.2). The results showed that the pI of construct A (pI = 7.79) was more appropriate for injection than construct B (pI = 5.08), because construct A is in the physiological fluid range. Both constructs also showed similar solubility, as was predicted by SolPro server. However, construct A also presented a partially better solubility than construct B. Afterward, the immunogenicity and allergenicity of both constructs were assessed, which were appropriate in both constructs (Table.2).

The GRAVY value for construct A was −0.353 compared to −0.397 of construct B, which indicates higher hydrophilicity of construct B than construct A. GRAVY value is the sum of hydropathic values of all amino acids in the input sequence divided by the number of the residues present in the sequence (Kyte and Doolittle, 1982). Lower values of GRAVY indicate more hydrophilicity possessed by the molecule.

Following the homology modeling of the vaccine constructs by Cluspro server, both models were evaluated by comparing the results of ProSA-web, ERRAT, and PROCHECK servers to identify the model with best quality. ProSA-Web results demonstrated that construct B had a more negative z-score, which indicated a better model quality than construct A. on the other hand, ERRAT and PROCHECK results showed more numbers of high probability residual errors in construct B compared to construct A, which supported the selection of construct A as the final vaccine construct with a better overall quality. The 3D models of both constructs were employed as the inputs for molecular docking, while TLR agonists should be pocketed inside the hydrophobic pocket of MD2 molecule, construct A proved to be the better overall model than construct B due to better pocketing inside the MD2 molecule.

Finally, the MD simulation was used to monitor the behavior of TLR4/MD2-vaccine complex over 90 ns. MD simulation results indicated that changes in RMSD and Rg plots have been mainly arisen from the changes of number of hydrogen bonds. When hydrogen bonds reached a steady state (almost from the 60th ns of the simulation time), the RMSD and Rg plots also became stable. It seems that hydrogen bonds played a critical role in the binding of the vaccine molecule to the MD2 protein. Number of hydrogen bonds between MD2 and TLR4 protein chains endured insignificant changes over the MD simulation time similar to their RMSD and Rg plots. It seems that there were a few displacements in the hydrogen bonds, but the number of hydrogen bonds in the primary and final structures of the MD simulation analysis were almost equal. Moreover, RMSF values of the vaccine molecule (Fig. 7.c) showed that the linker residues between RpfB at the N-terminal and the epitopes of L2 was the most flexible area of the vaccine molecule. The final structure obtained at the end of the MD simulation time showed the vaccine molecule was strongly bent at this area. Bending of the vaccine molecule was also happened on the linker residues between RpfB at the C-terminal and the epitopes of E6, although to a lesser extent. In fact, reducing the radius of gyration was mainly due to these bends. Some of the adjuvant residues located around the mentioned residues also tolerated high fluctuations. The residues present in the epitope parts showed almost low fluctuations, although some of them that were located near the linker residues underwent more fluctuations. Some of adjuvant residues interacting with the MD2 protein directly almost showed high fluctuations. These fluctuations were arisen from their efforts to obtain a better position in the MD2 cavity over the MD simulation period. Totally, the MD simulation results approved the stability of the designed vaccine binding to TLR4/MD2 complex. These results clarified that all three protein chains endeavored to reinforce their interactions with each other during the 90 ns MD simulation time in a way that TLR4/MD2-vaccine complex moved to a more stable position. The compatible interactions between the vaccine molecule and TLR4/MD2 complex guided this movement.

These results clarified that all three protein chains attempted to reinforce their interactions with each other during the 90 ns MD simulation time in a way that TLR4/MD2-vaccine complex moved to a more stable position. The compatible interactions between the vaccine molecule and TLR4/MD2 complex guided this movement. In conclusion, data achieved by MD simulations demonstrated the stability of vital protein-protein interactions in the designated time-frame, which could imply the appropriate vaccine design both physiochemically and structurally.

5. Conclusion

In this study, a novel multi-epitope vaccine construct with prophylactic and therapeutic features against HPV was designed employing several epitopes from L2, E6, and E7 proteins of HPV16 through various immunoinformatics tools. In addition to appropriate immunogenicity, the physiochemical properties of the vaccine construct appeared to be suitable for vaccine application Molecular docking and MD studies performed on the selected vaccine candidate demonstrated proper protein-protein interaction between the construct and TLR4, which implies the adjuvant functionality of the employed TLR4 agonist.

Vaccine design by computational tools proves to be a resourceful strategy in comparison to conventional vaccine design methods by frugality in time and resources. However, the proposed construct should be further validated regarding its physicochemical characters and also be evaluated in terms of efficacy and safety in future experiments.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This study was supported by Grant no. 13435 from the Research Council of Shiraz University of Medical Sciences, Shiraz University of Medical Sciences, Shiraz, Iran.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.meegid.2021.105084.

References

Arai, R., Wriggers, W., Nishikawa, Y., Nagamune, T., Fujisawa, T., 2004. Conformations of variably linked chimeric proteins evaluated by synchrotron X-ray small-angle scattering. Proteins: Struct. Funct. Bioinform. 57 (4), 829–838.

Argos, P., 1990. An investigation of oligopeptides linking domains in protein tertiary structures and possible candidates for general gene fusion. J. Mol. Biol. 211 (4), 943–958.

Billod, J.-M., Lacetera, A., Guzmán-Caldentey, J., Martín-Santamaría, S., 2016. Computational approaches to toll-like receptor 4 modulation. Molecules. 21 (8), 994.

Boehr, D.D., Nussinov, R., Wright, P.E., 2009. The role of dynamic conformational ensembles in biomolecular recognition. Nat. Chem. Biol. 5 (11), 789–796.

Bruni, L., Diaz, M., Castellsagué, M., Ferrer, E., Bosch, F.X., de Sanjosé, S., 2010. Cervical human papillomavirus prevalence in 5 continents: meta-analysis of 1 million women with normal cytological findings. J. Infect. Dis. 202 (12), 1789–1799.

Campo, M.S., Roden, R.B., 2010. Papillomavirus prophylactic vaccines: established successes, new approaches. J. Virol. 84 (3), 1214–1220.

Cargnelutti, D.E., Sanchez, M.V., Alvarez, P., Boado, L.A., Mattion, N.M., Scodeller, E., 2013. Enhancement of Th1 Immune Responses to Recombinant Influenza Nucleoprotein by Ribi Adjuvant.
Carpenter, S., O'Neill, L.A., 2009. Recent insights into the structure of toll-like receptors and post-translational modifications of their associated signalling proteins. Biochem. J. 422 (1), 1–10.
Chen, X., Zaro, J.L., Shen, W.-C., 2013. Fusion protein linkers: property, design and functionality. Adv. Drug Deliv. Rev. 65 (10), 1357–1369.
Choi, H.G., Kim, W.S., Back, Y.W., Kim, H., Kwon, K.W., Kim, J.S., et al., 2015. Mycobacterium tuberculosis RpfE promotes simultaneous Th1-and Th17-type T-cell immunity via TLR4-dependent maturation of dendritic cells. Eur. J. Immunol. 45 (7), 1957–1971.
Colovos, C., Yeates TO, 1993. Verification of protein structures: patterns of nonbonded atomic interactions. Protein Sci. 2 (9), 1511–1519.
de Oliveira, L.M.F., Morale, M.G., Chaves, A.A.M., Cavalher, A.M., Lopes, A.S., de Oliveira, Diniz M., et al., 2015. Design, immune responses and anti-tumor potential of an HPV16 E6E7 multi-epitope vaccine. PLoS One 10 (9), e0138686.
Dimitrov, I., Bangov, I., Flower, D.R., Doytchinova, I., 2014. AllerTOP v. 2—a server for in silico prediction of allergens. J. Mol. Model. 20 (6), 2278.
Doorbar, J., Egawa, N., Griffin, H., Kranjec, C., Murakami, I., 2015. Human papillomavirus molecular biology and disease association. Rev. Med. Virol. 25, 2–23.
Doytchinova, I.A., Flower, D.R., 2007. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. BMC Bioinform. 8 (1), 4.
Faridgohar, M., Nikoueinejad, H., 2017. New findings of toll-like receptors involved in Mycobacterium tuberculosis infection. Pathog. Glob. Health. 111 (5), 256–264.
Ferlay, J., Soerjomataram, I., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., et al., 2015. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int. J. Cancer 136 (5). E359-E86.
Gasteiger, E., Hoogland, C., Gattiker, A., Wilkins, M.R., Appel, R.D., Bairoch, A., 2005. Protein identification and Analysis Tools on the ExPASy Server. The Proteomics Protocols Handbook. Springer, pp. 571–607.
Giannini, S.L., Hanon, E., Moris, P., Van Mechelen, M., Morel, S., Dessy, F., et al., 2006. Enhanced humoral and memory B cellular immunity using HPV16/18 L1 VLP vaccine formulated with the MPL/aluminium salt combination (AS04) compared to aluminium salt only. Vaccine. 24 (33–34), 5937–5949.
Hansson, T., Oostenbrink, C., van Gunsteren, W., 2002. Molecular dynamics simulations. Curr. Opin. Struct. Biol. 12 (2), 190–196.
Harper, D.M., 2009. Currently approved prophylactic HPV vaccines. Expert Rev. Vaccines. 8 (12), 1663–1679.
Harper, D.M., Demars, L.R., 2014. Primary strategies for HPV infection and cervical cancer prevention. Clin. Obstet. Gynecol. 57 (2), 256–278.
Hasan, Y., Furtado, L., Tergas, A., Lee, N., Brooks, R., McCall, A., et al., 2020. A Phase 1 trial assessing the safety and tolerability of a therapeutic DNA vaccination against HPV16 and HPV18 E6/E7 oncogenes after chemoradiation for cervical cancer. Int. J. Radiat. Oncol. Biol. Phys. 107 (3), 487–498.
Hegde, N.R., Gauthami, S., Sampath Kumar, H., Bayry, J., 2018. The use of databases, data mining and immunoinformatics in vaccinology: where are we? Expert Opin. Drug Discovery 13 (2), 117–130.
Huleatt, J.W., Jacobs, A.R., Tang, J., Desai, P., Kopp, E.B., Huang, Y., et al., 2007. Vaccination with recombinant fusion proteins incorporating toll-like receptor ligands induces rapid cellular and humoral immunity. Vaccine. 25 (4), 763–775.
Jagu, S., Malandro, N., Kwak, K., Yuan, H., Schlegel, R., Palmer, K.E., et al., 2011. A multimeric L2 vaccine for prevention of animal papillomavirus infections. Virology. 420 (1), 43–50.
Jagu, S., Kwak, K., Karanam, B., Huh, W.K., Damotharan, V., Chivukula, S.V., et al., 2013. Optimization of multimeric human papillomavirus L2 vaccines. PLoS One 8 (1), e55538.
Jiang, R.T., Schellenbacher, C., Chackerian, B., Roden, R.B., 2016. Progress and prospects for L2-based human papillomavirus vaccines. Expert Rev. Vaccines. 15 (7), 853–862.
Kalnin, K., Tibbitts, T., Yan, Y., Stegalkina, S., Shen, L., Costa, V., et al., 2014. Low doses of flagellin-L2 multimer vaccines protect against challenge with diverse papillomavirus genotypes. Vaccine. 32 (28), 3540–3547.
Kawana, K., Kawana, Y., Yoshikawa, H., Taketani, Y., Yoshiike, K., Kanda, T., 2001. Nasal immunization of mice with peptide having a cross-neutralization epitope on minor capsid protein L2 of human papillomavirus type 16 elicit systemic and mucosal antibodies. Vaccine. 19 (11−12), 1496–1502.
Kenter, G.G., Welters, M.J., Valentijn, A.R.P., Löwik, M.J., Berends-van der Meer, D.M., Vloon, A.P., et al., 2008. Phase I immunotherapeutic trial with long peptides spanning the E6 and E7 sequences of high-risk human papillomavirus 16 in end-stage cervical cancer patients shows low toxicity and robust immunogenicity. Clin. Cancer Res. 14 (1), 169–177.
Kim, Y., Ponomarenko, J., Zhu, Z., Tamang, D., Wang, P., Greenbaum, J., et al., 2012. Immune epitope database analysis resource. Nucleic Acids Res. 40 (W1). W525-W30.
Kim, J.S., Kim, W.S., Choi, H.G., Jang, B., Lee, K., Park, J.H., et al., 2013. Mycobacterium tuberculosis RpfB drives Th1-type T cell immunity via a TLR4-dependent activation of dendritic cells. J. Leukoc. Biol. 94 (4), 733–749.
Ko, J., Park, H., Seok, C., 2012a. GalaxyTBM: template-based modeling by building a reliable core and refining unreliable local regions. BMC Bioinform. 13 (1), 198.
Ko, J., Park, H., Heo, L., Seok, C., 2012b. GalaxyWEB server for protein structure prediction and refinement. Nucleic Acids Res. 40 (W1). W294-W7.
Kozakov, D., Hall, D.R., Xia, B., Porter, K.A., Padhorny, D., Yueh, C., et al., 2017. The ClusPro web server for protein–protein docking. Nat. Protoc. 12 (2), 255.
Kringelum, J.V., Lundegaard, C., Lund, O., Nielsen, M., 2012. Reliable B cell epitope predictions: impacts of method development and improved benchmarking. PLoS Comput. Biol. 8 (12), e1002829.
Krüger, C.L., Zeuner, M.-T., Cottrell, G.S., Widera, D., Heilemann, M., 2017. Quantitative single-molecule imaging of TLR4 reveals ligand-specific receptor dimerization. Sci. Signal. 10 (503).
Kumar, S., Sunagar, R., Gosselin, E., 2019. Bacterial protein toll-like-receptor agonists: a novel perspective on vaccine adjuvants. Front. Immunol. 10, 1144.
Kyte, J., Doolittle, R.F., 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157 (1), 105–132.
Larsen, J.E.P., Lund, O., Nielsen, M., 2006. Improved method for predicting linear B-cell epitopes. Immunome Res. 2 (1), 1–7.
Laskowski, R.A., MacArthur, M.W., Moss, D.S., Thornton, J.M., 1993. Procheck: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. 26 (2), 283–291.
Lata, S., Bhasin, M., Raghava, G.P., 2007. Application of Machine Learning Techniques in Predicting MHC Binders. Immunoinformatics. Springer, pp. 201–215.
Livingston, B., Crimi, C., Newman, M., Higashimoto, Y., Appella, E., Sidney, J., et al., 2002. A rational strategy to design multiepitope immunogens based on multiple Th lymphocyte epitopes. J. Immunol. 168 (11), 5499–5506.
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 Res. 36 (suppl_2). W509-W12.
Magnan, C.N., Randall, A., Baldi, P., 2009. SOLpro: accurate sequence-based prediction of protein solubility. Bioinformatics. 25 (17), 2200–2207.
Magnan, C.N., Zeller, M., Kayala, M.A., Vigil, A., Randall, A., Felgner, P.L., et al., 2010. High-throughput prediction of protein antigenicity using protein microarray data. Bioinformatics. 26 (23), 2936–2943.
María, R., Arturo, C., Alicia, J.A., Paulina, M., Gerardo, A.O., 2017. The Impact of Bioinformatics on Vaccine Design and Development. Vaccines: InTech, Rijeka, Croatia.
Mosaheb, M.M., Reiser, M.L., Wetzler, L.M., 2017. Toll-like receptor ligand-based vaccine adjuvants require intact MyD88 signaling in antigen-presenting cells for germinal center formation and antibody production. Front. Immunol. 8, 225.
Moyle, P.M., 2017. Biotechnology approaches to produce potent, self-adjuvanting antigen-adjuvant fusion protein subunit vaccines. Biotechnol. Adv. 35 (3), 375–389.
Negahdaripour, M., Eslami, M., Nezafat, N., Hajighahramani, N., Ghoshoon, M.B., Shoolian, E., et al., 2017. A novel HPV prophylactic peptide vaccine, designed by immunoinformatics and structural vaccinology approaches. Infect. Genet. Evol. 54, 402–416.
Negahdaripour, M., Nezafat, N., Eslami, M., Ghoshoon, M.B., Shoolian, E., Najafipour, S., et al., 2018. Structural vaccinology considerations for in silico designing of a multi-epitope vaccine. Infect. Genet. Evol. 58, 96–109.
Negahdaripour, M., Nezafat, N., Heidari, R., Erfani, N., Hajighahramani, N., Ghoshoon, M.B., et al., 2020. Production and preliminary in vivo evaluations of a novel in silico-designed L2-based potential HPV vaccine. Curr. Pharm. Biotechnol. 21 (4), 316–324.
Nielsen, M., Lundegaard, C., Worning, P., Lauemøller, S.L., Lamberth, K., Buus, S., et al., 2003. Reliable prediction of T-cell epitopes using neural networks with novel sequence representations. Protein Sci. 12 (5), 1007–1017.
Panatto, D., Amicizia, D., Bragazzi, N.L., Rizzitelli, E., Tramalloni, D., Valle, I., et al., 2015. Human Papillomavirus Vaccine: State of the Art and Future Perspectives. Advances in Protein Chemistry and Structural Biology, 101. Elsevier, pp. 231–322.
Park, B.S., Song, D.H., Kim, H.M., Choi, B.-S., Lee, H., Lee, J.-O., 2009. The structural basis of lipopolysaccharide recognition by the TLR4–MD-2 complex. Nature. 458 (7242), 1191–1195.
Peters, B., Sette, A., 2005. Generating quantitative models describing the sequence specificity of biological processes with the stabilized matrix method. BMC Bioinform. 6 (1), 132.
Reed, S.G., Hsu, F.-C., Carter, D., Orr, M.T., 2016. The science of vaccine adjuvants: advances in TLR4 ligand adjuvants. Curr. Opin. Immunol. 41, 85–90.
Saha, S., GPS, Raghava (Eds.), 2004. BcePred: Prediction of Continuous B-cell Epitopes in Antigenic Sequences using Physico-Chemical Properties. International Conference on Artificial Immune Systems. Springer.
Schellenbacher, C., Kwak, K., Fink, D., Shafti-Keramat, S., Huber, B., Jindra, C., et al., 2013. Efficacy of RG1-VLP vaccination against infections with genital and cutaneous human papillomaviruses. J. Investig. Dermatol. 133 (12), 2706–2713.
Schellenbacher, C., Roden, R.B., Kirnbauer, R., 2017. Developments in L2-based human papillomavirus (HPV) vaccines. Virus Res. 231, 166–175.
Schiller, J.T., Lowy, D.R., 2015. Raising expectations for subunit vaccine. J. Infect. Dis. 211 (9), 1373–1375.
Severson, J., Evans, T.Y., Lee, P., Chan, T.-S., Arany, I., Tyring, S.K., 2001. Human papillomavirus infections: epidemiology, pathogenesis, and therapy. J. Cutan. Med. Surg. Incorpor. Med. Surg. Dermatol. 5 (1), 43–60.
Sidney, J., Assarsson, E., Moore, C., Ngo, S., Pinilla, C., Sette, A., et al., 2008. Quantitative peptide binding motifs for 19 human and mouse MHC class I molecules derived using positional scanning combinatorial peptide libraries. Immunome Res. 4 (1), 2.
Soria-Guerra, R.E., Nieto-Gomez, R., Govea-Alonso, D.O., Rosales-Mendoza, S., 2015. An overview of bioinformatics tools for epitope prediction: implications on vaccine development. J. Biomed. Inform. 53, 405–414.
Torre, L.A., Bray, F., Siegel, R.L., Ferlay, J., Lortet-Tieulent, J., Jemal, A., 2015. Global cancer statistics, 2012. CA Cancer J. Clin. 65 (2), 87–108.
Tyler, M., Tumban, E., Chackerian, B., 2014. Second-generation prophylactic HPV vaccines: successes and challenges. Expert Rev. Vaccines. 13 (2), 247–255.
UniProt, 2021. The universal protein knowledgebase in 2021. Nucleic Acids Res. 49 (D1). D480-D9.

Vakili, B., Eslami, M., Hatam, G.R., Zare, B., Erfani, N., Nezafat, N., et al., 2018. Immunoinformatics-aided design of a potential multi-epitope peptide vaccine against Leishmania infantum. Int. J. Biol. Macromol. 120, 1127–1139.
Vambutas, A., DeVoti, J., Nouri, M., Drijfhout, J., Lipford, G., Bonagura, V., et al., 2005. Therapeutic vaccination with papillomavirus E6 and E7 long peptides results in the control of both established virus-induced lesions and latently infected sites in a pre-clinical cottontail rabbit papillomavirus model. Vaccine. 23 (45), 5271–5280.
Van der Burg, S., Kwappenberg, K., O'neill, T., Brandt, R., Melief, C., Hickling, J., et al., 2001. Pre-clinical safety and efficacy of TA-CIN, a recombinant HPV16 L2E6E7 fusion protein vaccine, in homologous and heterologous prime-boost regimens. Vaccine. 19 (27), 3652–3660.
van Leeuwen, H.C., Strating, M.J., Rensen, M., de Laat, W., van der Vliet, P.C., 1997. Linker length and composition influence the flexibility of Oct-1 DNA binding. EMBO J. 16 (8), 2043–2053.
van Poelgeest, M.I., Welters, M.J., Vermeij, R., Stynenbosch, L.F., Loof, N.M., Berends-van der Meer, D.M., et al., 2016. Vaccination against oncoproteins of HPV16 for noninvasive vulvar/vaginal lesions: lesion clearance is related to the strength of the T-cell response. Clin. Cancer Res. 22 (10), 2342–2350.
Wang, J.W., C-f, Hung, Huh, W.K., Trimble, C.L., Roden, R.B., 2015. Immunoprevention of human papillomavirus–associated malignancies. Cancer Prev. Res. 8 (2), 95–104.
Wiederstein, M., Sippl, M.J., 2007. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res. 35 (suppl_2). W407-W10.
Wu, C.-Y., Monie, A., Pang, X., Hung, C.-F., Wu, T., 2010. Improving therapeutic HPV peptide-based vaccine potency by enhancing CD4+ T help and dendritic cell activation. J. Biomed. Sci. 17 (1), 1–10.
Zajac, A.J., Murali-Krishna, K., Blattman, J.N., Ahmed, R., 1998. Therapeutic vaccination against chronic viral infection: the importance of cooperation between CD4+ and CD8+ T cells. Curr. Opin. Immunol. 10 (4), 444–449.
Zhai, L., Tumban, E., 2016. Gardasil-9: a global survey of projected efficacy. Antivir. Res. 130, 101–109.
Zhao, X., Yang, F., Mariz, F., Osen, W., Bolchi, A., Ottonello, S., et al., 2020. Combined prophylactic and therapeutic immune responses against human papillomaviruses induced by a thioredoxin-based L2-E7 nanoparticle vaccine. PLoS Pathog. 16 (9), e1008827.
Zur, Hausen H., 2002. Papillomaviruses and cancer: from basic studies to clinical application. Nat. Rev. Cancer 2 (5), 342–350.