Current Proteomics
Year 2020
ISSN: 1875-6247 (Online)
ISSN: 1570-1646 (Print)

In silico approach for designing a novel recombinant fusion protein as a Candidate Vaccine against HPV

Mohsen Sisakht[1], Amir Mahmoodzadeh[2], Mohammadsaeid Zahedi[3], Davood Rostamzadeh[4, 5], Amin Moradi Hasan-Abad[6], Amir Atapour[7,*]

[1]Department of Biochemistry, Shiraz University of Medical Sciences, Shiraz, Iran; [2]Medical Biology Research Center, Health Technology Institute, Kermanshah University of Medical Sciences, Kermanshah, Iran; [3]School of Pharmacy, Shiraz University of Medical Sciences, Shiraz, Iran; [4]Medicinal plants research center, Yasuj University of Medical Sciences, Yasuj, Iran; [5]Cellular and Molecular Research Center, Yasuj University of Medical Sciences, Yasuj, Iran; [6]Department of Medical Biotechnology, School of Advanced Technologies in Medicine, Tehran University of Medical Sciences, Tehran, Iran; [7]Department of Medical Biotechnology, School of Advanced Medical Sciences and Technologies, Shiraz University of Medical Sciences, Shiraz, Iran

Abstract: ***Background*:** Human papillomavirus (HPV) is the main biological agent causing sexually transmitted diseases (STDs), including precancerous lesions and several types of prevalent cancers. To date, numerous types of vaccines are designed to prevent high-risk HPV. However, their prophylactic effect is not the same and does not clear previous infections. Therefore, there is an urgent need for developing therapeutic vaccines that trigger cell-mediated immune responses for the treatment of HPV. The HPV16 E6 and E7 proteins are ideal targets for vaccine therapy against HPV. Fusion protein vaccines, which include both immunogenic interest protein and an adjuvant for augmenting the immunogenicity effects, are theoretically capable of guarantee the power of the immune system against HPV.

***Method*:** A vaccine construct, including HPV16 E6/E7 proteins along with a heat shock protein GP96 (E6/E7-NTGP96 construct), was designed using in silico methods. By the aid of the SWISS-MODEL server, the optimal 3D model of the designed vaccine was selected, followed by physicochemical and molecular parameters were performed using bioinformatics tools. Docking studies were done to evaluate the binding interaction of the vaccine. Allergenicity, immunogenicity, B, and T cell epitopes of the designed construct were predicted.

***Results*:** Immunological and structural computational results illustrated that our designed construct is potentially proper for stimulation of cellular and humoral immune responses against HPV.

***Conclusion*:** Computational studies showed that the E6/E7-NTGP96 construct is a promising candidate vaccine that needs further in vitro and in vivo evaluations.

*Address correspondence to this author at the Department of Medical Biotechnology, Faculty of Medical Biotechnology, School of Advanced Medical Sciences and Technologies, Shiraz University of Medical Sciences, Shiraz, 71348-14336 Iran; Tel: +98713 2305471; E-mail: amir.atapoor58@yahoo.com

Keywords: Human papillomavirus (HPV), *In silico*, Cancer, Vaccine, GP96, E6.

1. INTRODUCTION

Human Papillomavirus (HPV) is the main pathogen for different types of cancer, such as cervical, vulvar, penile, rectal, anal, and oropharyngeal [1-4]. Statistically, 530000 cases of cervical cancer patients have been exposed to HPV annually, in addition to an approximate number of 113,400 people who are suffering from other types of HPV-driven cancers [5, 6]. Since these cancers are known as preventable diseases, vaccination has been regarded as an effective preventing approach due to having several advantages such as cost-effectiveness, easy application, and the elimination of time-consuming, exhausting, and stressful processes of treatment. Nowadays, several vaccines against HPV infection, such as Gecolin and Cecolin, are available in the market [7]. However, some challenges still exist to face in vaccine development, such as covering all medically relevant genotypes, the political and anthropological vaccine acceptance, the shelf life, and safety concerns. Recently, numerous antigens and adjuvants were used to be applied in vaccine development that enhances the magnitude quality and longevity of specific immune responses to antigens and with having minimum toxicity [8, 9]. There are more than 100 genetically different HPVs, which all of them have an 8kb double-stranded DNA genome that encoding six early genes (E1-E7) that are needed for viral replication and two late genes (L1 & L2) that are essential for virion assembly [10]. At least 14 HPVs are known as carcinogenic [11, 12], and HPV-16 and HPV-18 are solely the cause of about 80% of cervical cancers and precancerous cervical lesions [12]. Hence, licensed HPV vaccines are generally designed to target HPV-16 and HPV-18 [13]. Accumulated data have shown that viral E6 and E7 proteins are the main contributors to the development of HPV-induced cervical cancer, which are required to preserve the malignant phenotype [14-17].

Adjuvants are essential components of vaccines that increase the magnitude and duration of immune responses [18]. Numerous licensed adjuvants could be used for human vaccines, such as endogenous immunomodulators like bacterial cell wall components, cytokines, and Toll-like receptors (TLRs) [19, 20]. In previous studies, heat shock proteins (HSPs) have been reported as effective adjuvant molecules. Although the ability of HSPs in antigen presentation and cross-presentation has remained controversial, there is evidence indicating that HSPs could bind and present tumor-associated antigens and finally activate anti-tumor CD4+ and CD8+ T cells [21]. The immunogenicity of HSP preparations has been successfully attributed to HSPs-fused peptides [22]. The heat shock protein GP96, an endoplasmic reticulum chaperone, is a substantial chaperon for the regulation of both innate and adaptive immunity [23]. GP96 could be used as an adjuvant and is able to be transferred, along with peptides, to the MHC class I molecules of APC cells. Besides, GP96 may play the same role as a cytokine in the activation of dendritic cells (DCs) and proliferation of CD8+ T cells [24-26]. However, the N-terminal fragment of GP96 up-regulates the same costimulatory receptors and could induce secretion of the same cytokines [27, 28], which makes it a promising adjuvant solely. Various studies have been shown that the N-terminal domain of GP96 (NTGP96) has the same adjuvant properties. For instance, it has been reported that the fusion of the N-terminal domain of GP96 with Hepatitis-B surface antigen (HBsAg) or hepatitis C virus antigens could act as a potent adjuvant for enhancing immunity and create stronger immunization in treated animals [29-32].

Developing a new vaccine is a time-intensive and costly process; therefore, we applied bioinformatics approaches to design a new vaccine against HPV. In this research, we tried to design a novel fusion protein composed of E6/E7 and NTGP96 proteins (E6/E7-NTGP96), and after that, we investigated the physicochemical properties of E6/E7-NTGP96 fusion protein using bioinformatics tools. Then, after secondary and three dimensional (3D) structures prediction, the best 3D model of the fusion protein was selected for interaction with TLR2 and TLR4 receptors. Also, an immunoanalysis such as B, T cell epitopes prediction, and allergenicity, immunogenicity were performed. All in all, in silico evaluations, indicated that our fusion protein (E6 / E7-NTGP96) probably represents a potential candidate vaccine against HPV and could be considered a suitable vaccine in the development process. However, in vitro and in vivo tests should confirm this claim in the future.

2. MATERIALS AND METHOD

2.1. Structural evaluations

2.1.1. Sequence Design and Construction

The protein sequences of HPV 16 E6 (PDB ID: P03126), HPV 16 E7 (PDB ID: P03129), Heat Shock Protein, GP96 (PDB ID: Q7ZTB3) as our adjuvant, human TLR2 (PDB ID: 2Z81) and TLR4 (PDB ID: 3VQ2) sequences were obtained from UniProt database at http://www.uniprot.org. The E6/E7-NTGP96 construct was built using eight glycines (GGGGGGGG) as a linker to join the N- terminal part of GP96 sequences with the HPV 16 E6/E7 protein sequences.

2.1.2. Physicochemical Properties Analysis

All of the physiochemical properties were defined using the Expasy ProtParam online server at http://web.expasy.org/protparam/. Indices as such Grand Average of Hydropathicity (GRAVY) and molecular weight (MW), in vivo and in vitro half-life, instability index, extinction coefficient, and theoretical isoelectric point were all investigated during the process [33, 34].

2.1.3. Secondary Structure Prediction

We used the GOR4 tool to predict the secondary structure of E6/E7-NTGP96 fusion protein. GOR4 method is based on information theory and the assumption that the information function of a protein chain could be approximated by a sum of information from single residues and pairs of residues [35].

2.1.4. Protein Solubility Prediction

To evaluate the fusion protein solubility, we used the SOLpro server at http://scratch.proteomics.ics.uci.edu. This server predicts the protein solubility under overexpression circumstances in *Escherichia coli*. It applies a two-stage SVM method due to the multiple representations of the primary sequence [36].

2.1.5. The 3D Model Designing of Vaccine Construct

Several servers, such as GALAXY web, I-TASSER, SWISS-MODEL, and Phyre2, were utilized to achieve 3D best structure models [37-40]. SWISS-MODEL 3D models earned the highest score in validation results; therefore, its models were selected to pursue the protein construction process. Also, SWISS-MODEL is an online server (issmodel.expasy.org) that is initially designed for protein structure modeling. SWISS-MODEL applies numerous tools to determine the validity of the model. The similarity to B-factors in crystal structures and the related column in SWISS-MODEL result files consists of a C-score, giving an estimate of the variability of the template structures at this position are among them. Finally, a detailed log file containing all of the stages endowed to the user [37].

2.1.6. Validation of the 3D model

Potential errors are a hidden part of every kind of simulation; thus, it is indispensable for them to be analyzed and being refined again. For this reason, the validation assessment was done through ProSA-web at https://prosa.services.came.sbg.ac.at/prosa.php. Z-score which shows the overall model quality and calculates the deviation of the total energy regarding an energy distribution coming from random conformations [41], ERRAT server at http://servicesn.mbi.ucla.edu/ERRAT/ in which chooses between correctly and incorrectly determined regions protein structures according to characteristic atomic interactions, using statistical methods that differentiate errors in model building, gives rise to increase randomized distributions of the different atom types from correct distributions [42] and Ramachandran plot analysis at http://mordred.bioc.cam.ac.uk/~rapper/rampage.php by making a Ramachandran plot, it could be determined which torsional angles are permitted and could obtain insight into the structure of peptides.

2.1.7. Refinement of 3D model

The best structure created by the SWISS-MODEL online server was introduced to the Galaxy web server for the refinement process. This tool firstly rebuilds side chains and repacks side-chain, followed by overall structure relaxation using molecular dynamics simulation [43].

2.1.8. Fused-Protein Docking with TLR2 & TLR4

To evaluate the interaction between our fusion protein and TLR2 and TLR4 as potential receptors, we used the SwarmDock server at https://bmm.crick.ac.uk/~svc-bmm-swarmdock/. It uses an algorithm with a normal modes approach in which docks with the aid of a hybrid particle swarm optimization [44].

2.1.9. Docked model Refinement

Here, to refine the best-selected ligand/receptor complex structure produced by docking analysis, the FireDock online tool at http://bioinfo3d.cs.tau.ac.il/FireDock/php.php tool was used. The FireDock server is the first web server for flexible refinement and scoring of protein-protein docking solutions. It includes optimization of side-chain conformations and rigid-body orientation and permits a high-throughput refinement. The method refines each candidate and ranks all the candidates according to the global energy [45].

2.1.10. In silico codon adaptation and cloning of the vaccine construct

Codon adaptation is a method for analyzing expressing the foreign genes in the host when the codon usage is different in both organisms. In this study, the Java Codon Adaptation Tool (JCAT) (http://www.jcat.de/) was used for codon adaptation analysis. The Java Codon Adaptation Tool (JCAT) presents a simple method to adapt the Codon Usage to most sequenced prokaryotic organisms and selected eukaryotic organisms. Therefore, the designed vaccine sequence was submitted for expressing in *E. coli* K12. Expression rate parameters such as the codon adaptation index (CAI), GC-content, and other additional options contain avoid rho-independent transcription terminators, prokaryotic ribosome binding sites, and cleavage sites of restriction enzymes were selected.

In the final construct sequence, the restriction sites NcoI and XhoI were included at the N and C-terminal sites, respectively. Eventually, the Vector NTI Advance v10.0 software was applied to clone of the adapted sequence of the designed vaccine construct into the pET28a (+) vector.

2.2. Immunoinformatic Evaluations

2.2.1. Allergenicity and Antigenicity Prediction

AllerTOP 1.0 and AlgPred servers were used to survey allergenicity. AllerTOP 1.0 is Python-based, with a GUI written in HTML at http://www.ddg-pharmfac.net/AllerTOP/, comprises of 2210 allergens and 2210 non-allergens from the same species by classification of the *k*-nearest neighbor algorithm (*k*NN), whereas *k*=3 and 5-fold cross-validated. AlgPred webserver http://www.imtech.res.in/raghava/algpred/ was also used for the evaluation of allergenicity. It has six different approaches; therefore, we took the most sensitive and specific one: The hybrid approach. This approach combines SVM and IgE epitope-based method results in better sensitivity along with higher specificity [46, 47]. VaxiJen v2.0 and ANTIGENpro servers evaluate the antigenicity of the final fusion protein. The Vaxijen v2.0 is at http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html, the first server for alignment-free prediction of protective antigens, designed to classify antigens based on the physicochemical characteristics of proteins disregarding the recourse to sequence alignment. The accuracy of the server ranges from 70 to 89%, considering the default threshold and target organisms, such as viral, bacteria, and tumor protein datasets [48]. The ANTIGENpro at http://scratch.proteomics.ics.uci.edu/ obtains protein antigenicity microarray data to predict protein antigenicity. It is an alignment-independent, sequenced based, and pathogen-unrelated approach [49].

2.2.2. The Prediction of B-cell

BCPRED was also used as a server (http://ailab.ist. psu.edu/bcpred/predict.html) for the prediction of the B-cell epitope server because it compares predictions of several methods, and consensus predictions are more reliable than individual predictions, the BCPREDS server also permits users to use each of the BCPred or AAP methods.

CBTOPE method at (http://crdd. osdd.net/raghava/cbtope/) predicts the conformational B cell epitopes using primary antigen sequence, disregarding any homology with the known structures.

ElliPro server was used for predicting antibody epitopes, which implements three algorithms performing the presenting facts: Firstly, delicate estimation of the protein shape as an ellipsoid; second, calculation of the residue protrusion index (PI) and clustering of neighboring residues due to their PI values (Ponomarenko et al., 2008).

Another implemented server in this study was Discotope at (http://www.cbs.dtu.dk/services/DiscoTope/), which predicts discontinuous B-cell epitopes from 3D protein structures [50].

2.2.3. CTL Epitope Prediction

CTLPred server at (http://crdd.osdd.net/raghava/ctlpred/) was used for the prediction of CTL epitopes. This approach is based on elegant machine learning techniques, such as an artificial neural network (ANN) and support vector machine (SVM). Besides, This method benefits techniques such as quantitative matrices based, consensus approach, and combined approach [51]. In this study, the combined approach was used, which has 75.8% accuracy.

2.2.4. Interferon-gamma inducing epitopes prediction

The identification of interferon-gamma inducing epitopes was also studied to achieve maximum induction of the immune system. IFNepitope server at (http://crdd.osdd.net/raghava/ifnepitope/) was used for the prediction of potential IFN-gamma inciting epitopes. This server applies several approaches such as machine learning technique, motifs-based search, and hybrid method according to an IFN-γ inducing and non-inducing MHC class II binders dataset. It is notable to say that the hybrid model earned the highest score due to its accuracy (81.39%) [52].

2.2.5. MHC-I Binding Epitope Prediction

NetMHC 4.0 server (http://www.cbs.dtu.dk/services/ NetMHC/) were used to determine the binding epitopes of MHC-I. The server is using artificial neural networks (ANNs). Moreover, a sequence alignment method, namely IEDB-NN 4.0, was also applied to make insertions and deletions in the alignment. In this method, the top x% predicted peptides contain the peptide of choice, which is a strong binder if the peptide has a %rank between 0.5% (default threshold for the strong binders) and 2% (default threshold for the weak binders) it is regarded as a weak binder [53, 54].

2.2.6. MHC-II Binding Epitope Prediction

NetMHC 2.3 server (http://www.cbs. dtu.dk/services/NetMHCII/) was used to determine MHC-II binding epitopes. NetMHC II predicts peptide binding using ANNs method for human HLA-DR, HLA-DQ, HLA-DP, and mouse MHC class II alleles [54, 55].

3. RESULTS

3.1. The physicochemical properties

The E6 and E7 protein sequence of HPV type 16 was joined by GGGGGGGG linker to the N-terminal of GP96 (NTGP96) (Fig. **1**). The designed construct has 594 amino acids consists of 80 positively and 111 negatively charged residues. The results revealed that the MW and theoretical pI of protein was 67.3 kDa and 4.97, respectively. The instability index was 52.56, that shows the vaccine is relatively unstable (it is above 40), and the aliphatic index (74.69) demonstrates that the designed vaccine is stable in a broad range of temperatures. The estimated half-lives were 30 hours (mammalian reticulocytes, in vitro), > 20 hours (yeast, in vivo), and > 10 hours (*E. coli*, in vivo). Aliphatic index and GRAVY were defined 74.69 and – 0.717, respectively. The GRAVY index was – 0.713. The SOLpro server as a solubility prediction tool was 0.74% for our fusion protein upon overexpression in *E. coli*.

Fig. (1). The schematic illustration of the E6/E7-GP96 (NTGP96) fusion construct. GGGGGGGG linker was used to join two parts.

3.2. Secondary Structure Prediction

GOR4 was used to predict the secondary structure of the fusion protein (Fig. **2**). It showed that or protein comprises 44.31% random coil, 19.90% extended strand, and 35.79% alpha helix.

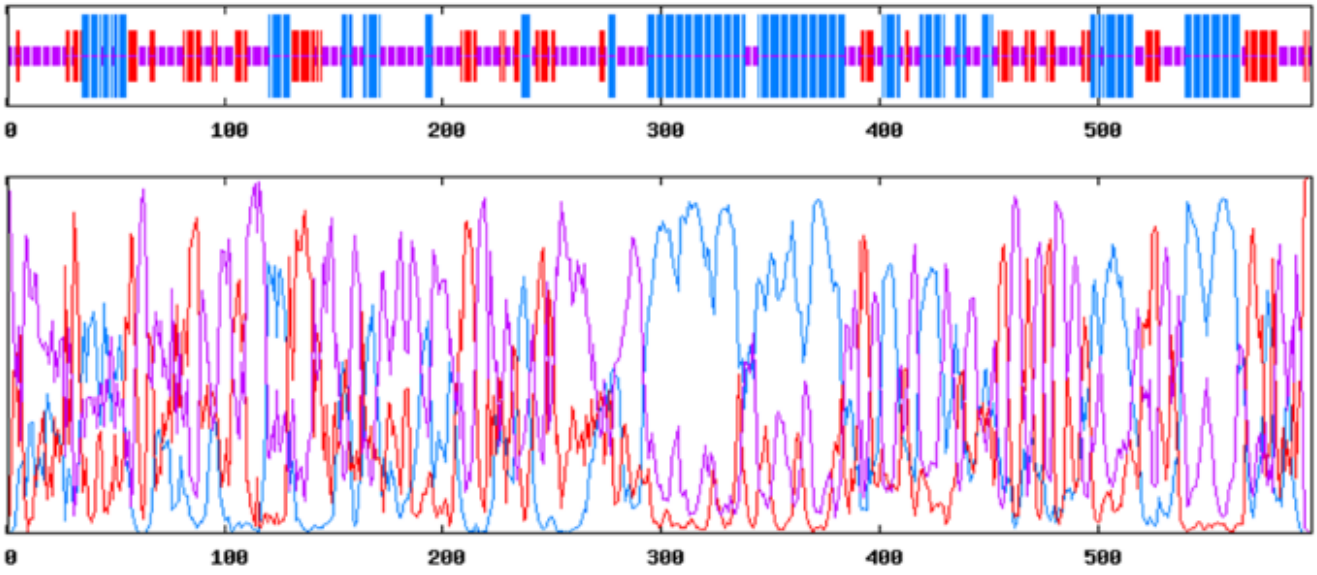

Fig. (2). Graphical illustration of the secondary structure of the fusion protein vaccine. Helix, extended strand and random coiled structures are represented in blue, purple, and red colors, respectively.

3.3. 3D Structure Modeling, Validation, and Refinement

We used a GGGGGGGG linker to connect the NTGP96 (as an adjuvant) to the E6/E7 protein sequence. The 3D structure modeling of the protein construct was performed using the SWISS-MODEL server due to its betterment in results than I-

TASSER, phyre2, and GALAXY web servers. Among the primary models predicted by SWISS-MODEL software, the number 1 model was chosen for the later refinement process. The refinement process was done using the GalaxyRefine server to reach the best 3D model. The primary and refined models are demonstrated in Fig. **3**. The calculation of Prosa z-score of both primary and refined structures indicates the numbers of – 4.98, – 6.04, respectively (Fig. **4**). Besides, the ERRAT server computes the overall quality factors of primary and final models as 90.5579 and 92.391, respectively (Fig. **5**). To observe the validation of 3D built models of fusion protein more precisely, Ramachandran plots (Fig. **6**) was done on models before and after the refinement process. The percentages of in favored, allowed, and outlier regions for pre-refined protein were 87.7%, 9.6%, and 2.7%, respectively, while after the refinement process, these percentages turned to 97.6%, 2.4%, and 0.0% respectively.

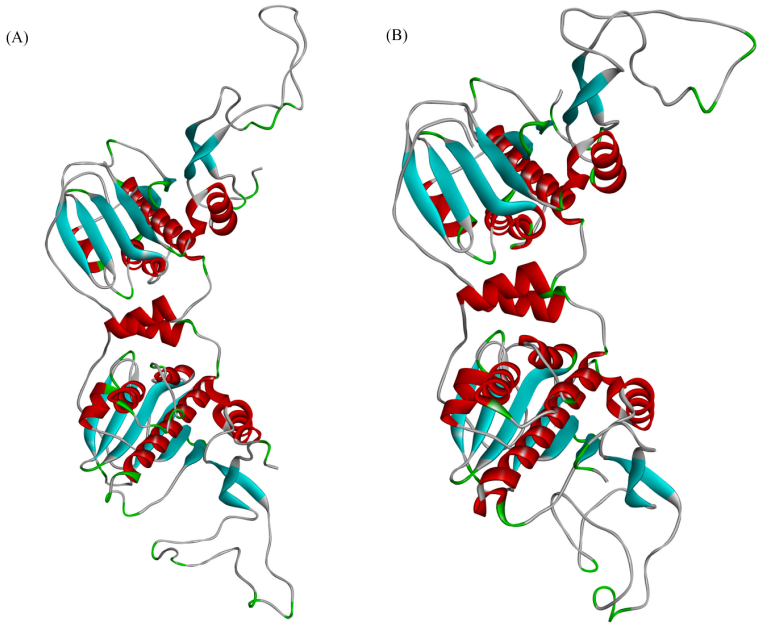

Fig. (3). The 3D structure of the designed construct.

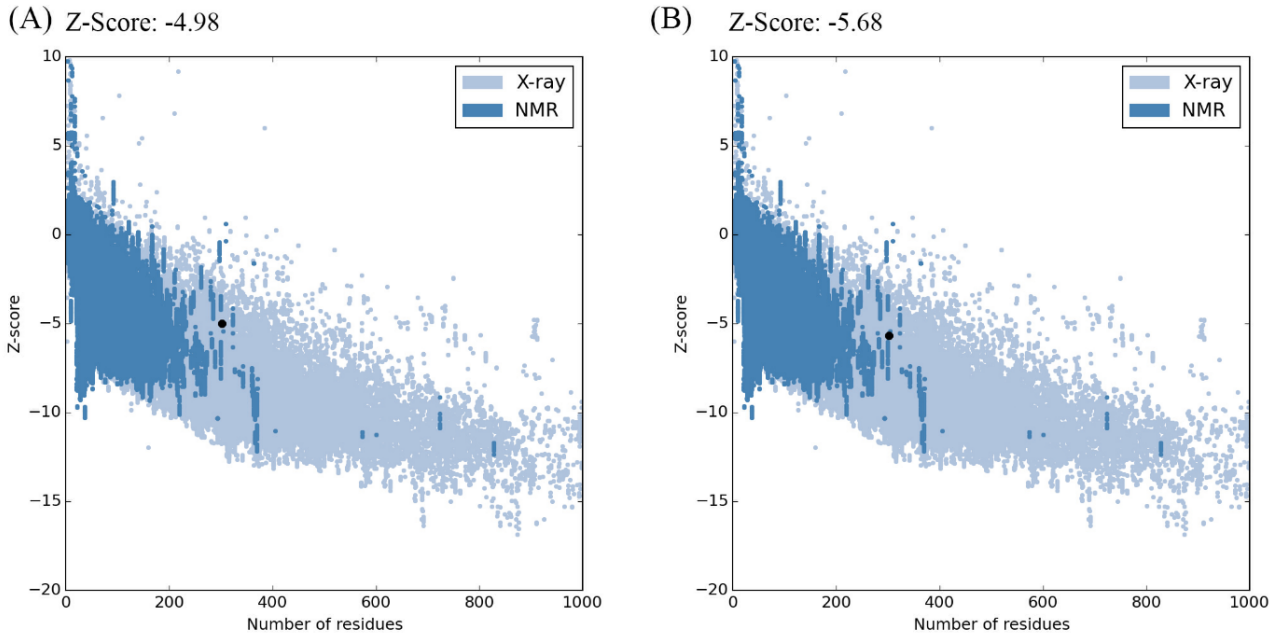

Fig. (4). ProSA Z-score plot. A. the Z-score for the primary model is: **-** 4.98 B. the overall quality factor for the final model is: **-** 6.04

Fig. (5). Errat plots. A. the overall quality factor for the primary model is 90.5579. B. the overall quality factor for the final model is 92.391.

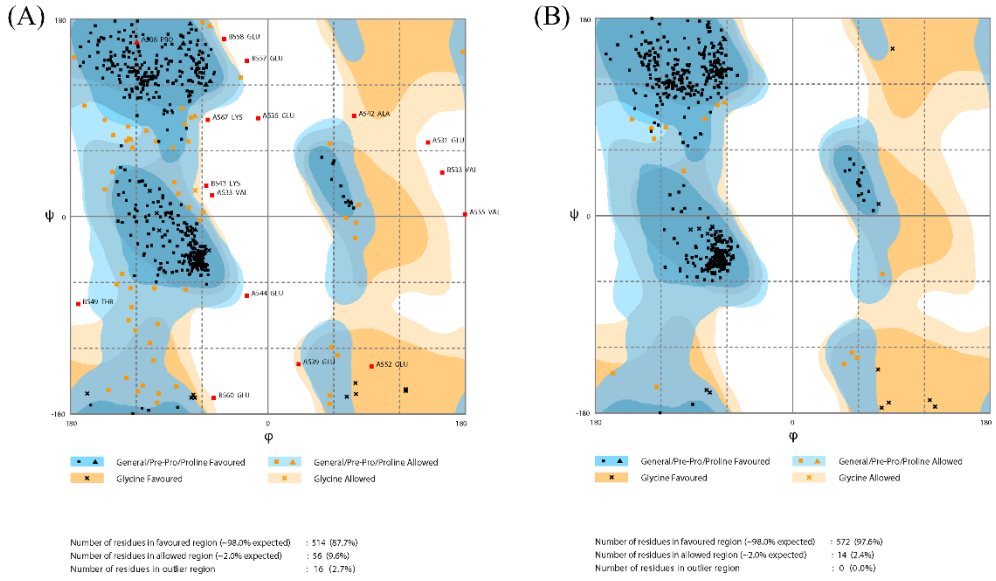

Fig. (6). Ramachandran plots of fusion construct. Initial model (A), refined model (B).

3.4. Docking Studies

The molecular docking of the 3D models was performed with SwarmDock to evaluate the interaction between the 3D designed vaccine model and TLR2 and TLR4 receptors (Fig. **7**). The three docked models were selected, and then the best model based on global energy was chosen by FireDock software.

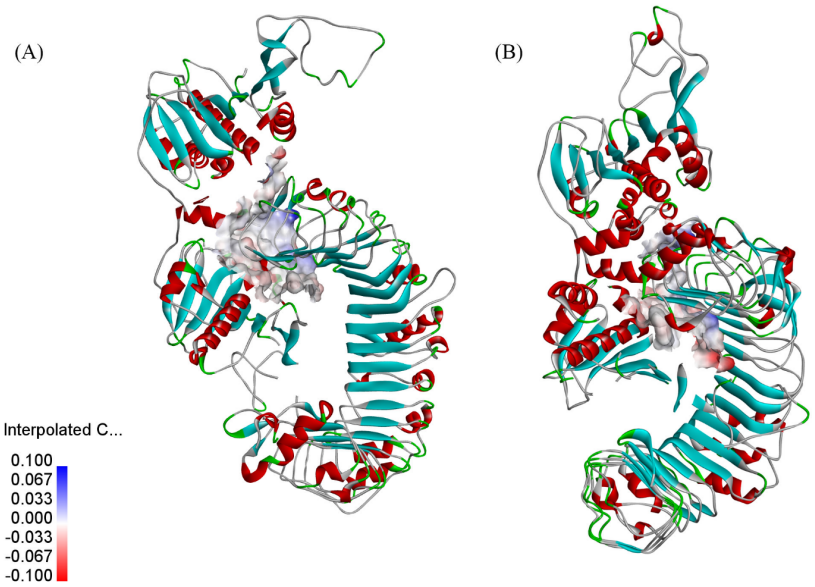

Fig. (7). The interaction of the fusion protein vaccine and TLR2 (A). The interaction of the fusion protein vaccine and TLR4 (B).

3.5. Docked model Refinement

The top three refined models of molecular docking processes based on global energy were – 47.8, – 45.1, and – 42.2 for the interaction of TLR2 and – 54.2, – 44.21, and – 42.3 for the interaction of TLR4, respectively.

3.6. Codon adaptation and in silico cloning of the vaccine construct

The Codon adaptation process, according to the *E. coli* K12 strain, was performed using the JCAT tool. In this study, we found that the GC-content and the value of the codon adaptive index (CAI) of the optimized sequence were 49.88% and 0.97, respectively. These indices were showed a good adaptation, which warranted the high rate of expression of the vaccine constructs in *E. coli* K12. Subsequently, the restriction sites of NcoI and XhoI were included, and the optimized vaccine sequence cloned into the vector pET28a (+), as shown in Fig. (8)**.** Also, for purification purposes, at the C-terminal end of the target sequence was used 6-histidine tag residues. The optimized codon sequence was 1794 nucleotides; thus, created a cloned vaccine of 7163 base pairs (Fig. **8**).

3.7. Antigenicity and Allergenicity Assessment

The predicted antigenicity for the fusion protein was 0.6880% by the Vaxijen server and 0.922481% by the ANTIGENpro server at a 0.4% threshold for the bacterial model. The evidence suggests that the built fusion protein may be regarded as a potential antigen. The amounts of allergenicity also were achieved using AllerTOP and AlgPred servers. AllerTOP considered our fusion protein a non-allergen.

3.8. Prediction of B-cell

The result of BCPRED, CBTOPE, ElliPro, and DiscoTope servers for the antibody epitopes prediction of our 590 fusion protein residues is shown in Table **1**, **2**, **3**, and **4**, respectively.

3.9. CTL Epitopes Prediction

The top ten of CTL epitopes in the protein vaccine obtained from the CTLpred server are listed in Table **5**.

3.10. IFN-Gamma inducing Epitopes Prediction

As to illustrate IFN-γ induced epitope results, top-ten of the IFN-γ inducers with the highest scores is listed in Table **6**. As it is shown, 5 of the final IFN-γ inducing epitopes have got a score above 1.

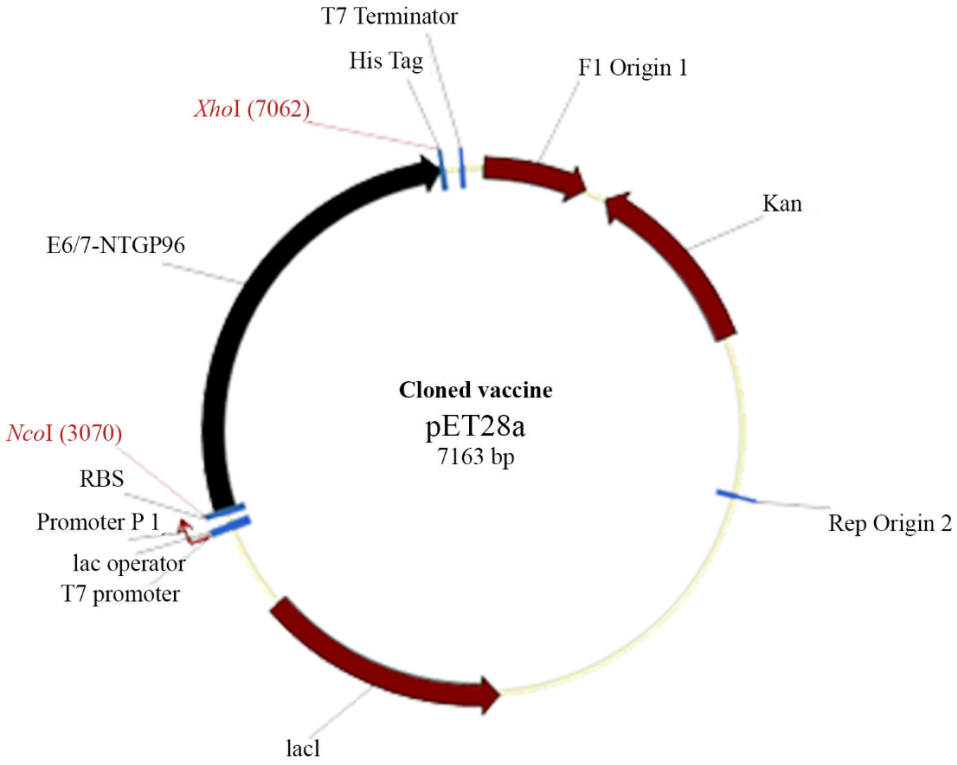

Fig. (8). In *silico* cloning of the optimized codon sequence of the fusion construct (Black color) into the *E. coli* pET28a (+) vector.

Table 1. BCPred Predictions of 20 amino acids epitope length (Prepared by BCPRED server 1.0).

Position	Epitope	Score
249	GPICSQKPGGGGGGGGMEDD	1
549	TDEEAAVEEEDEEKKPKTKK	1
188	DSSEEEDEIDGPAGQAEPDR	1

Table 2. The conformational B-cell epitopes predicted by the CBTOPE.

Position	Amino acid	Probability scale
66-73	PYAVCDKC	5
211-213	NIV	5
210	Y	6

Table 3. Predicted B-cell epitope residues of vaccine structure retrieved from ElliPro server.

No	Residues	Number Of Residues	Score
1	A:K545, A:D546, A:E547	3	0.915
2	A:E504, A:L505, A:E506, A:W526, A:S527, A:S528, A:K529, A:T530, A:E531, A:T532, A:V533, A:E534, A:E535, A:P536, A:L537, A:D538, A:E539, A:E540, A:E541, A:A542, A:K543, A:E544,	38	0.81

	A:D559, A:E560, A:K563, A:P564, A:K565, A:T566, A:K567, A:K568, A:V569, A:E570, A:K571, A:T572, A:V573, A:W574, A:D575, A:E577		
3	A:Q517, A:F518, A:N520, A:D548, A:T549, A:D550, A:E551, A:E552, A:A553, A:A554, A:V555, A:E556, A:E557, A:E558, A:E561, A:K562, A:L578, A:M579, A:N580, A:D581, A:I582, A:K583, A:P584, A:I585, A:W586, A:Q587, A:R588, A:P589, A:S590, A:K591, A:E592, A:I593, A:E594	33	0.77

Table 4. The conformational B-cell epitopes from 3D protein structure by the DiscoTope 2.0 server.

Amino acid	Residue number	Contact number	Propensity score	Discotop score
LYS	563	4	6.957	5.697
LYS	543	0	6.299	5.575
GLU	540	0	5.572	4.932

Table 5. CTLPred server for ten high ranked CTL epitopes.

Peptide Rank	Start position	Sequence	Score (ANN/SVM)	Prediction
1	418	GTSEFLSKL	0.97/1.2348071	Epitope
2	61	YRGGNPYAV	0.72/1.2738093	Epitope
3	298	SREEEAIQL	0.88/1.0674995	Epitope

Table 6. IFN-gamma induced epitope results.

Sequence	Method	Score
GPICSQKPGGGGGGGGMEDD	SVM	1.6599565
DWELMNDIKPIWQRPSKEIE	SVM	1.6534435
KISEYRHYCYSLYGTTLEQQ	SVM	1.424779

3.11. MHC peptide prediction

MHC I and II epitopes prediction are shown in Tables **7** and **8**, respectively.

Table 7. MHC-I binding epitopes using NetMHC 4.0 server * (Strong Binder (SB) threshold: <%Rank 0.500).

Position	HLA	Peptide Sequence	Binding Type	Affinity(nM)	%Rank
332	B0801	MMKLIINSL	SB*	9.02	0.01
367	A0101	LTDDQALAA	SB	13.74	0.03
60	B3901	YRGGNPYAV	SB	12.37	0.03

Table 8. MHC-II binding epitopes using NetMHC 2.3 server * (Strong Binder (SB) threshold: <%Rank 2).

Allele	Position	Peptide Sequence	Binding Type	Affinity(nM)	%Rank
DRB4_0101	362	KIRLMSLTDDQALAA	SB*	20.3	0.60
DRB4_0101	361	DKIRLMSLTDDQALA	SB	13.3	0.25
DRB4_0101	360	LDKIRLMSLTDDQAL	SB	10.6	0.1

4. DISCUSSION

During the last decade, the introduction of HPV 16/18 vaccines has led to a meaningful decrease in high-grade cervical lesions [56, 57]. In the efficient and pervasive vaccines era, an effective multi-potent vaccine which could be cost-benefitted is essential to confront HPV-driven cancers. However, in recent years, various studies have been conducted on an effective vaccine against HPV based on E6/E7 antigens.

Lee DW et al. investigated the adenoviral vaccine expressing E6 and E7 proteins as suppressors of the growth of the HPV-positive tumor [58]. Discoveries showed that the splenocytes of vaccinated mice along with Ad5-E6/E7 only produced E6/E7-specific IFN-γ. Fourteen days after vaccination, the E6/E7-specific IFN-γ production dramatically elevated until day 70.

Successful experiments were performed in binding GP96 to HPV16 E7 oncoprotein, generating DNA vaccine, and evaluating the vaccine in mice. Such as Daemi and et al. have discovered that E7 linkage with C-Terminal of GP96 DNA vaccine can hugely regress TC-1 tumor and augment the longevity (about 50 days) in the experimental group compared to control negative group. It is inferred from the mentioned study that structural domains of immune chaperones can stimulate an immune response against cancer [59].

The investigation of Bolhassani et al., which was done in the C57BL/6 mice model, showed that DNA vaccination, including E7 and Gp96, induced Th1 response. They concluded that the co-delivery of naked DNA E7 + Gp96 plasmid was immunologically more effective than E7 alone [22].

The interactions between GP96 and TLR2 and TLR4 and its potency to be used as an immunogen factor have been studied in several studies [60, 61]. In an in silico study, the E6/E7 proteins represented as potential vaccines, among the 20 clusters of CTL epitopes in E6, cluster 3 contained the most epitopes. Protein E7 clusters were also analyzed, and cluster 3 was with the most epitopes [62]. Negahdaripour *et al*. used HPV minor capsid protein (L2) along with flagellin adjuvant to trigger an optimum immune response, and it has appropriate results [63].

Therefore, in this study, we used E6/E7 proteins as HPV antigens along with an NTGP96 as an adjuvant for candidate vaccine construct.

A suitable linker should be applied, which can impact the folding rate, protein stability, and domain orientations to fuse different parts of the designed vaccine construct [64-67].

The studies have shown that linkers containing S and G residues can enhance the solubility as well as Proteolysis resistance, and also help the moieties to adjoin domains, act freely and to be accessible [68-70]. We used a linker of 8 glycines (GGGGGGGG) with high flexibility, to build the vaccine by its two moieties, the first one was E6/E7 proteins of HPV-16, and the other part was N-terminal of GP96, a kind of HSP 90 (E6/E7-NTGP96) (Fig. **1**).

In our study, the physicochemical properties of the designed construct (E6/E7-NTGP96) were retrieved via the ProtParam tool. It indicated that the total 80 number of positively charged residues versus the total 111 number of negatively charged residues, which implies that our fusion construct has a negative charge [33]. The next important factor is the extinction coefficient shows the amount of light absorption at a specific wavelength for a chemical substance. The extinction coefficient depends on Tyrosine (Y), Cysteine (C) and Tryptophan (W) as the content of amino acids. The extinction coefficient of our 67 kDa designed structure at 280nm was 58,090 $M^{-1}cm^{-1}$, which is the result of the high concentration of Tyrosine (Y), Cysteine (C) and Tryptophan (W) amino acids and thereby is an indicator of the susceptibility of UV analysis methods use. Our fusion protein has acidity property because the theoretical pI was 4.97. The theoretical pI method applies the isoelectric focusing approach to develop buffer systems for protein purification [71, 72]. The instability index presents an estimation of the stability of the protein in a test tube. A protein with an instability index smaller than 40 is predicted as a stable, a value above 40 predicts that the protein may be unstable [33]. The instability index of 53.57 classifies our protein as an unstable protein.

Nevertheless, the aliphatic index shows our fusion protein is stable in a broad range of temperatures. This index defines the occupancy rate of protein by Valine (V), Isoleucine (I), Leucine (L), and Alanine (A), as well as measures the thermostability of globular proteins [73, 74]. Another valuable thing to consider is the GRAVY index, which was – 0.713 in our case; it is shown that our fusion protein is hydrophilic nature and its high tendency to water versus oil. The secondary structure of proteins has a special effect on the 3D model and the function of proteins. Here, we used the PSIPRED server to predict the secondary structure. It indicated that the random coil structures were somewhat higher than other structures (44.31%), which may be due to high amounts of proline and glycine in our fusion protein (Fig. 2). Especially, proline residues cause disruption of the secondary structure and the occurrence of polypeptide chains torsion [75, 76], according to the influences of 3D protein models on its relevant function. Therefore, knowing the details of the three-dimensional structure of our fusion protein is beneficial in this regard [73, 74], we generated 3D models of the designed construct by four servers (Galaxy web, I-TASSER, SWISS-MODEL, and Phyre) (Fig. 3).

All models obtained were then evaluated by the ProSA-web, RAMPAGE, and ERRAT servers (Fig 4-6), to know the potential errors and improve the quality of the predicted 3D model. Therefore, according to the validation data, models obtained from the

SWISS-MODEL server were chosen for further evaluations. In the following, model 3 produced by the SWISS-MODEL server as the final model was selected, and the data showed that the final model needs to be refined. Here, we refined the final model via GalaxyRefine software. The best of refining model underwent molecular docking with TLR2 and TLR4 by SwarmDock server (Fig. 7). The Firedock server classified the models based on global energy level, then the least energy of models was chosen to reach the refined docking model (– 47.8 for the bond between TLR2 and -54.2 for the bond between TLR4). In this study, the codon optimization to improve the transcriptional and translational efficiency and analysis of the expression rate of our recombinant vaccine protein were performed in the E. *coli* K12 strain. The CAI value (0.97) closer to 1.0 that indicated the higher expression probability of our vaccine protein and GC content of the optimized construct sequence was 49.88%. Fortunately, the results were satisfying and imply a successful refinement process for our experiment. Bioinformatics data illustrated that the fusion protein has acceptable antigenicity and solubility, without any allergenicity, thus could be recommended as a potent vaccine against HPV. Also, immunoinformatics evaluation including B-cell (linear and conformational B-cell epitopes) (Tables 1-4), CTL (Table 5), and IFN-γ (Table 6) inducing epitopes were all revealed that the vaccine was capable of inducing the humoral and cellular immune system appropriately. In more detail, BCPreds and CBTOPE were used to evaluate linear and conformational B-cell epitopes (Table 1 and Table 2), while to define conformational B-cell epitopes, two servers of DiscoTope and Ellipro helped us to determine the epitopes (Table 3 and Table 4). CBTOPE is sequence-based, while DiscoTope is based on structure information. As for CTL and IFN-γ epitopes, the CTLPred server and IFNepitope server were applied to predict the CTL and IFNgamma epitopes, represented in Table 5 and Table 6, respectively. [50, 77, 78]. As previously mentioned in this article, MHC I and MHC II play an important role against HPV, therefore, we used NetMHC 4.0 and NetMHC 2.3 servers and thereby the results showed that MHC I alleles and MHC II alleles (Table 7, 8) were capable of HPV inhibition and elimination in silico environment.

CONCLUSIONS

A vaccine-based strategy to induce immune system responses is one of the most effective ways to fight HPV cancer. In this study, we designed a fusion protein and analyzed it using different bioinformatics tools. The results of the bioinformatics evaluations indicated that the designed vaccine construct could be a potential candidate vaccine against HPV cancer. However, in vitro and in vivo immunological tests are required to support this claim in the future.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable.

HUMAN AND ANIMAL RIGHTS

No Animals/Humans were used for studies that are the basis of this research.

CONSENT FOR PUBLICATION

Not applicable.

AVAILABILITY OF DATA AND MATERIALS

The authors confirm that the data supporting the findings of this research are available within the article.

FUNDING

None.

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

The authors wish to thank Shiraz University of Medical Sciences, Shiraz, Iran for supporting the conduct of this research.

REFERENCES

[1] Lin, C.; Franceschi, S.; Clifford, G.M. Human papillomavirus types from infection to cancer in the anus, according to sex and HIV status: a systematic review and meta-analysis. *Lancet Infect. Dis.,* **2018**, *18*(2), 198-206. http://dx.doi.org/10.1016/S1473-3099(17)30653-9 PMID: 29158102

[2] Al-Sahaf, S.; Hunter, K.D.; Bolt, R.; Ottewell, P.D.; Murdoch, C. The IL-1/IL-1R axis induces greater fibroblast-derived chemokine release in human papillomavirus-negative compared to positive oropharyngeal cancer. *Int. J. Cancer,* **2019**, *144*(2), 334-344. http://dx.doi.org/10.1002/ijc.31852 PMID: 30191960

[3] Roden, R.B.S.; Stern, P.L. Opportunities and challenges for human papillomavirus vaccination in cancer. *Nat. Rev. Cancer,* **2018**, *18*(4), 240-254. http://dx.doi.org/10.1038/nrc.2018.13 PMID: 29497146

[4] McGuire, S. *World cancer report 2014*; Geneva, Switzerland: World Health Organization, international agency for research on cancer, WHO Press, 2015. Oxford University Press, **2016**.

[5] Forman, D.; de Martel, C.; Lacey, C.J.; Soerjomataram, I.; Lortet-Tieulent, J.; Bruni, L.; Vignat, J.; Ferlay, J.; Bray, F.; Plummer, M.; Franceschi, S. Global burden of human papillomavirus and related diseases. *Vaccine,* **2012**, *30*(Suppl. 5), F12-F23.

http://dx.doi.org/10.1016/j.vaccine.2012.07.055 PMID: 23199955
[6] Plummer, M.; de Martel, C.; Vignat, J.; Ferlay, J.; Bray, F.; Franceschi, S. Global burden of cancers attributable to infections in 2012: a synthetic analysis. *Lancet Glob. Health,* **2016**, *4*(9), e609-e616.
http://dx.doi.org/10.1016/S2214-109X(16)30143-7 PMID: 27470177
[7] Clendinen, C.; Zhang, Y.; Warburton, R.N.; Light, D.W. Manufacturing costs of HPV vaccines for developing countries. *Vaccine,* **2016**, *34*(48), 5984-5989.
http://dx.doi.org/10.1016/j.vaccine.2016.09.042 PMID: 27771183
[8] Dubensky, T.W., Jr Reed, S. G.*Adjuvants for cancer vaccines, Seminars in immunology*; Elsevier, **2010**, pp. 155-161.
[9] Bolhassani, A.; Shahbazi, S.; Milani, A.; Nadji, S.A. Small Heat Shock Proteins B1 and B6: Which One is the Most Effective Adjuvant in Therapeutic HPV Vaccine? *IUBMB Life,* **2018**, *70*(10), 1002-1011.
http://dx.doi.org/10.1002/iub.1892 PMID: 30171788
[10] zur Hausen, H. Papillomaviruses and cancer: from basic studies to clinical application. *Nat. Rev. Cancer,* **2002**, *2*(5), 342-350.
http://dx.doi.org/10.1038/nrc798 PMID: 12044010
[11] Humans, I.W.G.E.C.R.t.; Meeting, I.W.G.E.C.R.H.; Organization, W.H.; Cancer, I.A.R.o. *Human papillomaviruses*; World Health Organization, **2007**, Vol. 90, .
[12] Geraets, D.; Alemany, L.; Guimera, N.; de Sanjose, S.; de Koning, M.; Molijn, A.; Jenkins, D.; Bosch, X.; Quint, W.; Group, R.H.T.S. RIS HPV TT Study Group. Detection of rare and possibly carcinogenic human papillomavirus genotypes as single infections in invasive cervical cancer. *J. Pathol.,* **2012**, *228*(4), 534-543.
http://dx.doi.org/10.1002/path.4065 PMID: 22711526
[13] Kalu, N.N.; Mazumdar, T.; Peng, S.; Shen, L.; Sambandam, V.; Rao, X.; Xi, Y.; Li, L.; Qi, Y.; Gleber-Netto, F.O.; Patel, A.; Wang, J.; Frederick, M.J.; Myers, J.N.; Pickering, C.R.; Johnson, F.M. Genomic characterization of human papillomavirus-positive and -negative human squamous cell cancer cell lines. *Oncotarget,* **2017**, *8*(49), 86369-86383.
http://dx.doi.org/10.18632/oncotarget.21174 PMID: 29156801
[14] Bubeník, J. Animal models for development of therapeutic HPV16 vaccines (review). *Int. J. Oncol.,* **2002**, *20*(1), 207-212.
http://dx.doi.org/10.3892/ijo.20.1.207 PMID: 11743666
[15] Kotecha, M.T.; Afghan, R.K.; Vasilikopoulou, E.; Wilson, E.; Marsh, P.; Kast, W.M.; Davies, D.H.; Caparros-Wanderley, W. Enhanced tumour growth after DNA vaccination against human papilloma virus E7 oncoprotein: evidence for tumour-induced immune deviation. *Vaccine,* **2003**, *21*(19-20), 2506-2515.
http://dx.doi.org/10.1016/S0264-410X(03)00043-4 PMID: 12744885
[16] Hoppe-Seyler, K.; Bossler, F.; Braun, J.A.; Herrmann, A.L.; Hoppe-Seyler, F. The HPV E6/E7 oncogenes: key factors for viral carcinogenesis and therapeutic targets. *Trends Microbiol.,* **2018**, *26*(2), 158-168.
http://dx.doi.org/10.1016/j.tim.2017.07.007 PMID: 28823569
[17] Martínez-Ramírez, I.; Carrillo-García, A.; Contreras-Paredes, A.; Ortiz-Sánchez, E.; Cruz-Gregorio, A.; Lizano, M. Regulation of Cellular Metabolism by High-Risk Human Papillomaviruses. *Int. J. Mol. Sci.,* **2018**, *19*(7)E1839
http://dx.doi.org/10.3390/ijms19071839 PMID: 29932118
[18] Lee, S.; Nguyen, M.T. Recent advances of vaccine adjuvants for infectious diseases. *Immune Netw.,* **2015**, *15*(2), 51-57.
http://dx.doi.org/10.4110/in.2015.15.2.51 PMID: 25922593
[19] Medzhitov, R. Toll-like receptors and innate immunity. *Nat. Rev. Immunol.,* **2001**, *1*(2), 135-145.
http://dx.doi.org/10.1038/35100529 PMID: 11905821
[20] Suresh, R.; Mosser, D.M. Pattern recognition receptors in innate immunity, host defense, and immunopathology. *Adv. Physiol. Educ.,* **2013**, *37*(4), 284-291.
http://dx.doi.org/10.1152/advan.00058.2013 PMID: 24292903
[21] Ciocca, D.R.; Cayado-Gutierrez, N.; Maccioni, M.; Cuello-Carrion, F.D. Heat shock proteins (HSPs) based anti-cancer vaccines. *Curr. Mol. Med.,* **2012**, *12*(9), 1183-1197.
http://dx.doi.org/10.2174/156652412803306684 PMID: 22804241
[22] Bolhassani, A.; Zahedifard, F.; Taghikhani, M.; Rafati, S. Enhanced immunogenicity of HPV16E7 accompanied by Gp96 as an adjuvant in two vaccination strategies. *Vaccine,* **2008**, *26*(26), 3362-3370.
http://dx.doi.org/10.1016/j.vaccine.2008.03.082 PMID: 18471945
[23] Schild, H.; Rammensee, H.-G. gp96—the immune system's Swiss army knife. *nature immunology,* **2000**.
[24] Singh-Jasuja, H.; Scherer, H.U.; Hilf, N.; Arnold-Schild, D.; Rammensee, H.G.; Toes, R.E.; Schild, H. The heat shock protein gp96 induces maturation of dendritic cells and down-regulation of its receptor. *Eur. J. Immunol.,* **2000**, *30*(8), 2211-2215.
http://dx.doi.org/10.1002/1521-4141(2000)30:8<2211::AID-IMMU2211>3.0.CO;2-0 PMID: 10940912
[25] Ramirez, S.R.; Singh-Jasuja, H.; Warger, T.; Braedel-Ruoff, S.; Hilf, N.; Wiemann, K.; Rammensee, H-G.; Schild, H. Glycoprotein 96-activated dendritic cells induce a CD8-biased T cell response. *Cell Stress Chaperones,* **2005**, *10*(3), 221-229.
http://dx.doi.org/10.1379/CSC-117R.1 PMID: 16184767
[26] Tamura, Y.; Yoneda, A.; Takei, N.; Sawada, K. Spatiotemporal regulation of Hsp90–ligand complex leads to immune activation. *Front. Immunol.,* **2016**, *7*, 201.
http://dx.doi.org/10.3389/fimmu.2016.00201 PMID: 27252703
[27] Vogen, S.; Gidalevitz, T.; Biswas, C.; Simen, B.B.; Stein, E.; Gulmen, F.; Argon, Y. Radicicol-sensitive peptide binding to the N-terminal portion of GRP94. *J. Biol. Chem.,* **2002**, *277*(43), 40742-40750.
http://dx.doi.org/10.1074/jbc.M205323200 PMID: 12189140
[28] Biswas, C.; Sriram, U.; Ciric, B.; Ostrovsky, O.; Gallucci, S.; Argon, Y. The N-terminal fragment of GRP94 is sufficient for peptide presentation via professional antigen-presenting cells. *Int. Immunol.,* **2006**, *18*(7), 1147-1157.
http://dx.doi.org/10.1093/intimm/dxl049 PMID: 16772370
[29] Pishraft-Sabet, L.; Kosinska, A.D.; Rafati, S.; Bolhassani, A.; Taheri, T.; Memarnejadian, A.; Alavian, S-M.; Roggendorf, M.; Samimi-Rad, K. Enhancement of HCV polytope DNA vaccine efficacy by fusion to an N-terminal fragment of heat shock protein gp96. *Arch. Virol.,* **2015**, *160*(1), 141-152.
http://dx.doi.org/10.1007/s00705-014-2243-8 PMID: 25348271
[30] Li, H-T.; Yan, J-B.; Li, J.; Zhou, M-H.; Zhu, X-D.; Zhang, Y-X.; Tien, P. Enhancement of humoral immune responses to HBsAg by heat shock protein gp96 and its N-terminal fragment in mice. *World J. Gastroenterol.,* **2005**, *11*(19), 2858-2863.
http://dx.doi.org/10.3748/wjg.v11.i19.2858 PMID: 15902719
[31] Li, H.; Zhou, M.; Han, J.; Zhu, X.; Dong, T.; Gao, G.F.; Tien, P. Generation of murine CTL by a hepatitis B virus-specific peptide and evaluation of the adjuvant effect of heat shock protein glycoprotein 96 and its terminal fragments. *J. Immunol.,* **2005**, *174*(1), 195-204.
http://dx.doi.org/10.4049/jimmunol.174.1.195 PMID: 15611241

[32] Yan, J.; Liu, X.; Wang, Y.; Jiang, X.; Liu, H.; Wang, M.; Zhu, X.; Wu, M.; Tien, P. Enhancing the potency of HBV DNA vaccines using fusion genes of HBV-specific antigens and the N-terminal fragment of gp96. *The Journal of Gene Medicine: A cross-disciplinary journal for research on the science of gene transfer and its clinical applications,* **2007**, *9*(2), 107-121.

[33] Gasteiger, E.; Hoogland, C.; Gattiker, A.; Wilkins, M.R.; Appel, R.D.; Bairoch, A. Protein identification and analysis tools on the ExPASy server.*The proteomics protocols handbook*; Springer, **2005**, pp. 571-607.
http://dx.doi.org/10.1385/1-59259-890-0:571

[34] Jahangiri, A.; Amani, J.; Halabian, R. In silico analyses of staphylococcal enterotoxin B as a DNA vaccine for cancer therapy. *Int. J. Pept. Res. Ther.,* **2018**, *24*(1), 131-142.
http://dx.doi.org/10.1007/s10989-017-9595-3

[35] Garnier, J.; Osguthorpe, D.J.; Robson, B. Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. *J. Mol. Biol.,* **1978**, *120*(1), 97-120.
http://dx.doi.org/10.1016/0022-2836(78)90297-8 PMID: 642007

[36] Magnan, C.N.; Randall, A.; Baldi, P. SOLpro: accurate sequence-based prediction of protein solubility. *Bioinformatics,* **2009**, *25*(17), 2200-2207.
http://dx.doi.org/10.1093/bioinformatics/btp386 PMID: 19549632

[37] Schwede, T.; Kopp, J.; Guex, N.; Peitsch, M.C. SWISS-MODEL: An automated protein homology-modeling server. *Nucleic Acids Res.,* **2003**, *31*(13), 3381-3385.
http://dx.doi.org/10.1093/nar/gkg520 PMID: 12824332

[38] Shin, W-H.; Lee, G.R.; Heo, L.; Lee, H.; Seok, C. Prediction of protein structure and interaction by GALAXY protein modeling programs. *Bio Design,* **2014**, *2*(1), 1-11.

[39] Kelley, L.A.; Sternberg, M.J. Protein structure prediction on the Web: a case study using the Phyre server. *Nat. Protoc.,* **2009**, *4*(3), 363-371.
PMID: 19247286

[40] Roy, A.; Kucukural, A.; Zhang, Y. I-TASSER: a unified platform for automated protein structure and function prediction. *Nat. Protoc.,* **2010**, *5*(4), 725-738.
PMID: 20360767

[41] Wiederstein, M.; Sippl, M. J. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic acids research,* **2007**.
http://dx.doi.org/10.1093/nar/gkm290

[42] Colovos, C.; Yeates, T.O. Verification of protein structures: patterns of nonbonded atomic interactions. *Protein Sci.,* **1993**, *2*(9), 1511-1519.
PMID: 8401235

[43] Heo, L.; Park, H.; Seok, C. GalaxyRefine: Protein structure refinement driven by side-chain repacking. *Nucleic Acids Res.,* **2013**, *41*(Web Server issue)W384-8
http://dx.doi.org/10.1093/nar/gkt458 PMID: 23737448

[44] Torchala, M.; Bates, P.A. Predicting the structure of protein–protein complexes using the SwarmDock web server.*Protein Structure Prediction*; Springer, **2014**, pp. 181-197.
http://dx.doi.org/10.1007/978-1-4939-0366-5_13

[45] Mashiach, E.; Schneidman-Duhovny, D.; Andrusier, N.; Nussinov, R.; Wolfson, H. J. FireDock: a web server for fast interaction refinement in molecular docking. *Nucleic acids research,* **2008**.
http://dx.doi.org/10.1093/nar/gkn186

[46] Dimitrov, I. Flower, D. R.; Doytchinova, I.*AllerTOP-a server for in silico prediction of allergens, BMC bioinformatics*; BioMed Central, **2013**, p. S4.

[47] Saha, S.; Raghava, G. AlgPred: prediction of allergenic proteins and mapping of IgE epitopes. *Nucleic acids research,* **2006**.
http://dx.doi.org/10.1093/nar/gkl343

[48] Doytchinova, I.A.; Flower, D.R. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinformatics,* **2007**, *8*(1), 4.
http://dx.doi.org/10.1186/1471-2105-8-4 PMID: 17207271

[49] Cheng, J.; Randall, A. Z.; Sweredoski, M. J.; Baldi, P. SCRATCH: a protein structure and structural feature prediction server. *Nucleic acids research,* **2005**.
http://dx.doi.org/10.1093/nar/gki396

[50] Haste Andersen, P.; Nielsen, M.; Lund, O. Prediction of residues in discontinuous B-cell epitopes using protein 3D structures. *Protein Sci.,* **2006**, *15*(11), 2558-2567.
http://dx.doi.org/10.1110/ps.062405906 PMID: 17001032

[51] Bhasin, M.; Raghava, G.P. Prediction of CTL epitopes using QM, SVM and ANN techniques. *Vaccine,* **2004**, *22*(23-24), 3195-3204.
http://dx.doi.org/10.1016/j.vaccine.2004.02.005 PMID: 15297074

[52] Dhanda, S.K.; Vir, P.; Raghava, G.P. Designing of interferon-gamma inducing MHC class-II binders. *Biol. Direct,* **2013**, *8*(1), 30.
http://dx.doi.org/10.1186/1745-6150-8-30 PMID: 24304645

[53] Andreatta, M.; Nielsen, M. Gapped sequence alignment using artificial neural networks: application to the MHC class I system. *Bioinformatics,* **2016**, *32*(4), 511-517.
http://dx.doi.org/10.1093/bioinformatics/btv639 PMID: 26515819

[54] Jensen, K.K.; Andreatta, M.; Marcatili, P.; Buus, S.; Greenbaum, J.A.; Yan, Z.; Sette, A.; Peters, B.; Nielsen, M. Improved methods for predicting peptide binding affinity to MHC class II molecules. *Immunology,* **2018**, *154*(3), 394-406.
http://dx.doi.org/10.1111/imm.12889 PMID: 29315598

[55] Nielsen, M.; Lundegaard, C.; Lund, O. Prediction of MHC class II binding affinity using SMM-align, a novel stabilization matrix alignment method. *BMC Bioinformatics,* **2007**, *8*(1), 238.
http://dx.doi.org/10.1186/1471-2105-8-238 PMID: 17608956

[56] Baldur-Felskov, B.; Dehlendorff, C.; Munk, C.; Kjaer, S.K. Early impact of human papillomavirus vaccination on cervical neoplasia--nationwide follow-up of young Danish women. *J. Natl. Cancer Inst.,* **2014**, *106*(3)djt460
http://dx.doi.org/10.1093/jnci/djt460 PMID: 24552678

[57] Saraiya, M.; Unger, E.R.; Thompson, T.D.; Lynch, C.F.; Hernandez, B.Y.; Lyu, C.W.; Steinau, M.; Watson, M.; Wilkinson, E.J.; Hopenhayn, C.; Copeland, G.; Cozen, W.; Peters, E.S.; Huang, Y.; Saber, M.S.; Altekruse, S.; Goodman, M.T.; Workgroup, H.T.C. HPV Typing of Cancers Workgroup. US assessment of HPV types in cancers: implications for current and 9-valent HPV vaccines. *J. Natl. Cancer Inst.,* **2015**, *107*(6)djv086
http://dx.doi.org/10.1093/jnci/djv086 PMID: 25925419

[58] Wu, J.; Chen, G.; Zhuang, F-C.; Gao, M.; Wu, C-D.; He, Z-L.; Jiang, Y-S.; Li, J-B.; Bao, J-Y.; Mao, Z-A. Long-term toxicity, pharmacokinetics and immune effects of a recombinant adenovirus vaccine expressing human papillomavirus 16 E6 and E7 proteins (HPV16 E6E7-Ad5 Vac) in primates. *Am. J. Transl. Res.,* **2018**, *10*(5), 1539-1551.
PMID: 29887967

[59] Daemi, A.; Bolhassani, A.; Rafati, S.; Zahedifard, F.; Hosseinzadeh, S.; Doustdari, F. Different domains of glycoprotein 96 influence HPV16 E7 DNA vaccine potency via electroporation mediated delivery in tumor mice model. *Immunol. Lett.,* **2012**, *148*(2), 117-125. http://dx.doi.org/10.1016/j.imlet.2012.10.003 PMID: 23085605

[60] Warger, T.; Hilf, N.; Rechtsteiner, G.; Haselmayer, P.; Carrick, D.M.; Jonuleit, H.; von Landenberg, P.; Rammensee, H-G.; Nicchitta, C.V.; Radsak, M.P.; Schild, H. Interaction of TLR2 and TLR4 ligands with the N-terminal domain of Gp96 amplifies innate and adaptive immune responses. *J. Biol. Chem.,* **2006**, *281*(32), 22545-22553. http://dx.doi.org/10.1074/jbc.M502900200 PMID: 16754684

[61] Cosin-Roger, J.; Spalinger, M.R.; Ruiz, P.A.; Stanzel, C.; Terhalle, A.; Wolfram, L.; Melhem, H.; Atrott, K.; Lang, S.; Frey-Wagner, I.; Fried, M.; Scharl, M.; Hausmann, M.; Rogler, G. Gp96 deficiency affects TLR4 functionality and impairs ERK and p38 phosphorylation. *PLoS One,* **2018**, *13*(2)e0193003 http://dx.doi.org/10.1371/journal.pone.0193003 PMID: 29447283

[62] Yao, Y.; Huang, W.; Yang, X.; Sun, W.; Liu, X.; Cun, W.; Ma, Y. 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,* **2013**, *31*(18), 2289-2294. http://dx.doi.org/10.1016/j.vaccine.2013.02.065 PMID: 23499609

[63] Negahdaripour, M.; Eslami, M.; Nezafat, N.; Hajighahramani, N.; Ghoshoon, M.B.; Shoolian, E.; Dehshahri, A.; Erfani, N.; Morowvat, M.H.; Ghasemi, Y. A novel HPV prophylactic peptide vaccine, designed by immunoinformatics and structural vaccinology approaches. *Infect. Genet. Evol.,* **2017**, *54*, 402-416. http://dx.doi.org/10.1016/j.meegid.2017.08.002 PMID: 28780192

[64] George, R.A.; Heringa, J. An analysis of protein domain linkers: their classification and role in protein folding. *Protein Eng.,* **2002**, *15*(11), 871-879. http://dx.doi.org/10.1093/protein/15.11.871 PMID: 12538906

[65] Robinson, C.R.; Sauer, R.T. Optimizing the stability of single-chain proteins by linker length and composition mutagenesis. *Proc. Natl. Acad. Sci. USA,* **1998**, *95*(11), 5929-5934. http://dx.doi.org/10.1073/pnas.95.11.5929 PMID: 9600894

[66] van Leeuwen, H.C.; Strating, M.J.; Rensen, M.; de Laat, W.; van der Vliet, P.C. Linker length and composition influence the flexibility of Oct-1 DNA binding. *EMBO J.,* **1997**, *16*(8), 2043-2053. http://dx.doi.org/10.1093/emboj/16.8.2043 PMID: 9155030

[67] Atapour, A.; Mokarram, P. MostafaviPour, Z.; Hosseini, S. Y.; Ghasemi, Y.; Mohammadi, S.; Nezafat, N. Designing a fusion protein vaccine against HCV: an in silico approach. *Int. J. Pept. Res. Ther.,* **2018**, •••, 1-12.

[68] Kavoosi, M.; Creagh, A.L.; Kilburn, D.G.; Haynes, C.A. Strategy for selecting and characterizing linker peptides for CBM9-tagged fusion proteins expressed in Escherichia coli. *Biotechnol. Bioeng.,* **2007**, *98*(3), 599-610. http://dx.doi.org/10.1002/bit.21396 PMID: 17394253

[69] Klement, M.; Liu, C.; Loo, B.L.W.; Choo, A.B-H.; Ow, D.S-W.; Lee, D-Y. Effect of linker flexibility and length on the functionality of a cytotoxic engineered antibody fragment. *J. Biotechnol.,* **2015**, *199*, 90-97. http://dx.doi.org/10.1016/j.jbiotec.2015.02.008 PMID: 25697559

[70] Lau, Y.H.; de Andrade, P.; McKenzie, G.J.; Venkitaraman, A.R.; Spring, D.R. Linear aliphatic dialkynes as alternative linkers for double-click stapling of p53-derived peptides. *ChemBioChem,* **2014**, *15*(18), 2680-2683. http://dx.doi.org/10.1002/cbic.201402374 PMID: 25354189

[71] Bjellqvist, B.; Hughes, G.J.; Pasquali, C.; Paquet, N.; Ravier, F.; Sanchez, J.C.; Frutiger, S.; Hochstrasser, D. The focusing positions of polypeptides in immobilized pH gradients can be predicted from their amino acid sequences. *Electrophoresis,* **1993**, *14*(10), 1023-1031. http://dx.doi.org/10.1002/elps.11501401163 PMID: 8125050

[72] Pihlasalo, S.; Auranen, L.; Hänninen, P.; Härmä, H. Method for estimation of protein isoelectric point. *Anal. Chem.,* **2012**, *84*(19), 8253-8258. http://dx.doi.org/10.1021/ac301569b PMID: 22946671

[73] Ashokan, K.; Pillai, M. In silico characterization of silk fibroin protein using computational tools and servers. *Asian Journal of Experimental Science,* **2008**.

[74] Ikai, A. Thermostability and aliphatic index of globular proteins. *J. Biochem.,* **1980**, *88*(6), 1895-1898. PMID: 7462208

[75] Sen Gupta, P.; Mandal, B.; Bandyopadhyay, A. In silico characterization of human cyclooxygenase using computational tools and servers. *Int J Institutional Pharmacy Life Sci,* **2013**, *3*, 111.

[76] Sivakumar, K.; Balaji, S. In silico characterization of antifreeze proteins using computational tools and servers. *J. Chem. Sci.,* **2007**, *119*(5), 571-579. http://dx.doi.org/10.1007/s12039-007-0072-y

[77] Ansari, H.R.; Raghava, G.P. Identification of conformational B-cell Epitopes in an antigen from its primary sequence. *Immunome Res.,* **2010**, *6*(1), 6. http://dx.doi.org/10.1186/1745-7580-6-6 PMID: 20961417

[78] Y.; Dobbs, D.; Honavar, V. Predicting linear B‐cell epitopes using string kernels. *Journal of Molecular Recognition: An Interdisciplinary Journal,* **2008**, *21*(4), 243-255. http://dx.doi.org/10.1002/jmr.893