Design of A Multi-Epitope Protein Vaccine Against Herpes Simplex Virus, Human Papillomavirus and Chlamydia Trachomatis as the Main Causes of Sexually Transmitted Diseases

Hesam Dorosti
Shiraz University of Medical Sciences
Sedigheh Eskandari
Shiraz University of Medical Sciences
Mahboubeh Zarei
Shiraz University of Medical Sciences
Navid Nezafat (n.nezafat@srbiau.ac.ir)
Shiraz University of Medical Sciences
Younes Ghasemi
Shiraz University of Medical Sciences

Research Article

Keywords: Epitope vaccine, Herpes simplex virus, human papillomavirus, Chlamydia trachomatis, Bioinformatics, Structural Vaccinology

Posted Date: December 29th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-133991/v1

Version of Record: A version of this preprint was published at Infection, Genetics and Evolution on November 1st, 2021. See the published version at https://doi.org/10.1016/j.meegid.2021.105136.

Design of a multi-epitope protein vaccine against Herpes simplex virus, Human papillomavirus and Chlamydia trachomatis as the main causes of sexually transmitted diseases

Hesam Dorosti [a,b] , Sedigheh Eskandari[a], Mahboubeh Zarei [a,b], Navid Nezafat [a,b*], Younes Ghasemi [a,b*]

[a] Pharmaceutical Sciences Research Center, Shiraz University of Medical Sciences, Shiraz, Iran.

[b] Department of Pharmaceutical Biotechnology, School of Pharmacy, Shiraz University of Medical Sciences, Shiraz, Iran.

Corresponding authors:

*Navid Nezafat: Pharmaceutical Sciences Research Center, Shiraz University of Medical Science, Shiraz, Iran

E-mail : n.nezafat@srbiau.ac.ir

*Younes Ghasemi, Department of Pharmaceutical Biotechnology, School of Pharmacy, Shiraz University of Medical Sciences, P.O. Box 71345-1583, Shiraz, Iran.

E-mail : ghasemiy@sums.ac.ir

Abstract

Background: Sexually transmitted diseases (STDs) have a profound effect on reproductivity and sexual health worldwide. According to world health organization (WHO) 375 million new case of STD, including *chlamydia trachomatis* (chlamydia), *Neisseria gonorrhoeae*, HSV, HPV has been reported in 2016. More than 30 diverse pathogenesis have identified to be transmitted through sexual intercourse. Of these, viral infections (hepatitis B, herpes simplex virus (HSV or herpes), HIV, and human papillomavirus (HPV) are incurable. However, symptoms caused by the incurable viral infections can be alleviated through treatment. Antimicrobial resistance (AMR) of sexually transmitted infections (STIs) to antibiotics has increased recent years, in this regard, vaccination is proposed as an important strategy for prevention or treatment of STDs. Vaccine against HPV 16 and 18 suggests a new approach for controlling STDs but until now, there is no prophylactic or therapeutic vaccine have been approved for HSV-2 and *Chlamydia trachomatis* (CT); in this reason, developing an efficient vaccine is inevitable. Recently, different combinatorial forms of subunit vaccines against two or three type of bacteria have been designed.

Results: In this study, to design a combinatorial vaccine against HSV, CT, and HPV, the E7 and L2 from HPV, glycoprotein D from HSV-2 and ompA from CT were selected as final antigens. Afterward, the immunodominant helper T lymphocytes (HTLs) and cytolytic T lymphocytes (CTLs) epitopes were chosen from aforesaid antigens. P30 (tetanus toxoid epitope) as universal T-helper were also added to the vaccine. Moreover, flagellin D_1/D_0 as TLR5 agonist and the RS09 as a TLR4 ligand were incorporated to N and C-terminals of peptide vaccine, respectively. Finally, all selected parts were fused together by appropriate linkers to enhance vaccine efficiency. The physicochemical, structural, and immunological properties of the designed vaccine protein were assessed. To achieve the best 3D model of the protein vaccine, modeling, refinement, and validation of modeled structures were also done. Docking evaluation demonstrated suitable interaction between the vaccine and TLR5. Moreover, molecular dynamics (MD) studies showed an appropriate and stable structure of protein and TLR5.

Conclusion: Based on immunoinformatic analysis, our vaccine candidate could potentially incite humoral and cellular immunities, which are critical for protection against HPV, HSV-2, and *chlamydia trachomatis*. It should be noted that, experimental studies are needed to confirm the efficacy of the designed vaccine.

Keywords: Epitope vaccine; Herpes simplex virus; human papillomavirus; Chlamydia trachomatis; Bioinformatics; Structural Vaccinology

Background

Sexually transmitted infections (STIs) possess a notable influence on reproductive and sexual health [1]. Despite available prevention strategies, more than one million STIs are infected worldwide each day.

Although over 30 pathogens have been known as the main cause of sexual diseases, eight of them are obviously related to the greatest amount of morbidity. Among them, chlamydia, gonorrhea, syphilis (bacterial STIs) and trichomoniasis (parasitic STI) are curable STIs, whereas viral pathogens such as, HIV, human papillomavirus (HPV), herpes simplex virus (HSV), and hepatitis B virus (HBV) can cause chronic or lifelong infections.

HSV and HPV are considered as life-threatening STIs. HPV type 16 and 18 are responsible for about 50% and 20% of cervical cancer, respectively, with 500000 new cases diagnosed annually cause 270000 deaths. The prevalence of HSV was estimated 491.5 million people for HSV-2 and more than 122-192 million people for genital HSV-1 in 2016 [2].

On the other hand, Chlamydia remains a major global health concern, which infected 127 million men and women aged 15-49 in 2016, according to the WHO [3]. The high global burden imposed on the healthcare system by STIs including reproductive, maternal-child health outcomes, such as genital symptoms, pregnancy complications, cancer, infertility, and enhanced HIV transmission.

[4, 5]. Concerns about the current screening programs highlight the complexities of controlling STI and the need for further research on designing effective STI vaccines [6].

Different scientific groups and companies are attempting to design an effective HSV vaccine, but so far there is no approved vaccine on the market [7, 8]. Although several therapeutic HSV-2 vaccine candidates such as HerpV and GEN-003/MM2 are in phase I/II clinical trials. In the context of HPV vaccine, presently, the two prophylactic licensed vaccine against HPV 16 and 18 (Cervarix and Gardasil) are used in many countries [9]. Controlling Chlamydia as the major priority of global health requires an efficient vaccine. To date, no vaccine has licensed for Chlamydia [10].

Recently, several subunit vaccine candidates have been developed against two or three kinds of bacterial and viral STIs; for instance, in 2013 a research group developed a multi-epitope DNA vaccine harboring HSV, HPV and HIV antigens that could elicit efficient humoral and cellular responses [11].

Various kinds of protective antigens are applied for designing a vaccine. The cellular and capsid proteins, including E7, E6, L2 and L1 are the most important antigens that are employed in HPV vaccine candidates [12, 13]. To design HSV-2 vaccine, envelope glycoproteins (gD, gB, gH, gL and gC) have received more attention than other antigens. On the other hand, major outer membrane protein (MOMP) has been introduced as a potent antigen in the context of chlamydia vaccine design.

Besides protective antigens, adjuvants are the important part of vaccines, especially in epitope-based protein vaccines. Recently, vaccinologists have focused on natural adjuvants such as toll like receptors (TLRs) ligands. Each TLR can detect specific ligand, e.g., TLR4, TLR2/6 and TLR1/2 recognize lipids and lipoprotein, and lipopolysaccharides (LPS); the flagellin and profilin are recognized by TLR5, TLR11, respectively. TLR1, 2, 3, 5 and 6 are dominantly found in the epithelia of female reproductive tract [14, 15]. In this reason, the ligands of above-mentioned TLRs can be used in designing STI vaccines.

In our project, first, protective antigens of HSV-2, HPV and *Chlamydia trachomatis* were retrieved from vaccine investigation and online information network (VIOLIN) server. E7 and L2 from HPV, glycoprotein D from HSV-2 and ompA from chlamydia were selected as final antigens. The immunodominant HTL and CTL epitopes were chosen from aforesaid antigens. P30 (tetanus toxoid epitope) as universal T-helper that bound to different MHC-II alleles was also added to the vaccine. Moreover, flagellin D_1/D_0 as TLR5 agonist and the APPHALS motif as a TLR4 ligand were incorporated to N and C-terminals of peptide vaccine, respectively. In the second stage, all the selected epitopes and parts were fused together by suitable linkers to enhance vaccine efficiency.

Method

Sequence retrieval

In the first step, with the help of Protegen software, all the protective antigens form *Chlamydia trachomatis*, HSV-2 and HPV were retrieved.

The amino acid sequences of HPV L2 (Accession no. AAQ10718), HPV E7 (Accession no. AAQ10713), HSV-2 glycoprotein D (Accession no. AAB72102), Chlamydia ompA (Accession no. AAD18834), flagellin D_1/D_0 *Pseudomonas aeruginosa* (Accession no. AAG04481) were retrieved from the National Centre for Biotechnology Information (NCBI). Different amino acids and DNA sequences were trimmed by standalone software CLC protein Workbench 5.6 software.

Immunoinformatics analyses

T cell MHC class I and II restricted epitope identification

Four mouse MHC I alleles, H-2Dd, H-2Kb, H-2Kk, H-2Kd, and H-2Ld as well as two mouse MHC II alleles, H-2Ad, and H-2Ab were selected in the current study [16]. 15-mer MHC class II epitope were identified using Consensus method, combinatorial library, netMHCII-2.2, netMHCII-1.1, Sturniolo, NetMHCIIpan, and IEDB recommended method were all applied in IEDB server. The second server used to predict the MHC-II binding epitopes was the MHCpred with accuracy about 90% [17]. The result of MHCPred is based on the IC_{50} values. Finally the RANKPEP server was applied for predicting MHC-II peptide binders which used the position-specific scoring matrices (PSSM) method [18]. In RANKPEP server, the binding threshold was set 4-6% [18].

Identification of CTL epitopes and TCR-peptide/peptide-MHC interfaces

CTLpred, which operates based on machine learning techniques and quantitative matrix, was used for the prediction of epitopes with the accuracy of 75.8%. In addition, PAComplex was employed to determine homologous peptide antigens and peptide antigens of a query from experimental peptide databases and complete pathogen genome databases [19].

Linear and discontinuous B-cell epitope prediction

ABCpred and BcePREDS were used for linear B-cell epitopes prediction. ABCpred applies a neural network approach with an accuracy about 65% [20]. Bcepred predicts B-cell epitopes based on physicochemical properties or a combination of properties. This server uses a combination of the polarity, flexibility, and surface hydrophilicity properties at a threshold of 2.38 to predict epitopes with an accuracy of 58.7% [21]. Following the 3D modeling of the

protein, the DiscoTope 2.0 server was applied to predict the conformational B-cell epitopes. In this server the final score was calculated by two methods: a) contact numbers derived from surface accessibility, b) a novel epitope partiality amino acid score [22]. The default threshold, sensitivity, and specificity of the server were -3.7, 0.47, 0.75 respectively.

Evaluation of antigenicity and allergenicity

ANTIGENpro that applies a sequence-based and alignment-free method, was used to predict an antigenicity index. The accuracy of server is about 76% by cross-validation experiments [23]. Moreover, VaxiJen v2.0 server was used to predict whole protein antigenicity. This server is based on auto cross covariance (ACC) transformation of protein sequences into uniform vectors of principal amino acid properties. Depending on target organisms, the accuracy of server varies from 70% to 89% [24]. The AlgPred and AllerTOP servers were employed to predict protein allergenicity with high accuracy. The accuracy of AlgPred base on Hybrid prediction approach (SVM+IgEepitope+ARPs BLASTMAST) is about 85% at a threshold 0.4 [25]. The Allertope server using some machine learning methods such as logistic regression (LR), decision tree (DT), naïve Bayes (NB), random forest (RF), multilayer perceptron (MLP) and k nearest neighbours (kNN can predict allergens with 85.3% accuracy [26].

Structural analyses

Evaluation of the physicochemical parameters of the vaccine

Physicochemical properties of the protein vaccine construct, including *in-vivo* and *in-vitro* half-life, aliphatic index, grand average of hydropathicity and amino acid composition were determined by the ProtParam server [27, 28].

Solpro and Protein-sol servers were used for the evaluation of the solubility of the heterologous peptide vaccine in *E. coli* [29]. The Solpro is a sequence-based prediction server. The overall accuracy of SOLpro is about 74% with a threshold of 0.5. The accuracy of the predicted soluble proteins is actually higher (77.5%) than the accuracy of the predicted insoluble proteins (71.5%) [30]. Protein–Sol provides a fast sequence-based method for predicting protein solubility; the server predict protein solubility from sequence and the fold state stability of a protein from structure [29].

Homology modeling, refinement and validation of the modeled structure

I-TASSER was used for the automated prediction of 3D protein structure [31]. The server is based on iterative threading assembly simulations. The quality of the predicted models by I-TASSER is estimated by a confidence score (C-score). A model with a high confidence has a higher C-score. In addition to I-TASSER, GalaxyTBM server (template-based modeling) was used to generate the 3D structure. This server predicts protein structure from the sequence when experimental structures of homologous proteins are available as templates and refines loop or terminus regions by ab initio modeling. For the visualization and control of the 3D structure of protein the Discovery studio 3.5 and PyMOL programs were applied.

Galaxyloop and GalaxyRefine were used for refinement of the tertiary modeled structure [33]. The GalaxyRefine server employs mild and aggressive relaxation methods for refinement the whole protein that generates one, and four structures, respectively. The GalaxyRefine improves the initial models with high probability (>50%) [34]. Loop modeling was done by GalaxyLoop server that employs an ab initio protein loop modeling method based on a conformational space annealing (CSA) technique that is a global energy optimization method[34]. Then, the refined 3D structures were validated by the ERRAT, RAMPAGE, verify 3D, and ProSA-web servers to recognize the potential errors in the models and compare the predicted 3D structures [35].

Molecular dynamic (MD) simulation of the multi-epitope vaccine

MD simulation was carried out by GROMACS 5.1.2 package with Amber ff99SB*-ILDN force filed. In so doing, the vaccine molecule was solvated with TIP3P model of water in a cubic box with at least 10 Å distance between the closest portion of the vaccine molecule, and the box edges. Pdb2gmx was used to generate protein topology. The steepest descent method (5000 steps) was used to remove unfavorable contacts in the system and minimize the initial energy of the system. Genion was used to neutralize the overall charges of the system through adding 0.15 M NaCl. NVT V-rescale ensemble was used at 500 ps to stabilize the temperature of system up to 300 K. Then, 1ns MD simulation in NPT ensemble (P = 1 atm) was used to relax the system. All bonds were constrained using the LINCS algorithm, and the exactitude of the algorithm was adjusted to 4 and 1 for lincs-order and lincs-iter, respectively. The Verlet scheme was used for the neighbor searching, and the short-range van der Waals cutoff and the short-range electrostatic cutoff were set to 1 nm. The main MD simulation was run for 50 ns. All analyses were calculated

by the gmx hbond, gmx rms, gmx sasa, and gmx rmsf tools of the GROMACS simulation package, version 5.1.2.

Docking analysis of the vaccine candidate with TLR5

ClusPro server was used for docking simulation between TLR5 (PDB ID: 3J0A) and vaccine [36]. Moreover, the online database PDB sum was utilized for schematic illustration of the interactions between vaccine and receptor [37]. Another server for the evaluation of protein-protein docking was SwarmDock, which uses a combination of particle swarm optimization and local docking to determine low-energy positions and orientations of the binding partners. Additionally, GalaxyRefineComplex tool at GalaxyWEB server was utilized for refinement of docked complex that enhances model accuracy in terms of both inter-protein orientation and interfaces contact through refining low-resolution docking structures. This method captures conformational changes occurring upon binding and improves flexibility in the overall docking structure and the protein interface. For symmetric homo-complexes, symmetric refinement is used.

Results

Selection of protective antigens

According to result of Protegen server five protective antigens were identified from HSV-2, HPV, and one (ompA) was defined from *Chlamydia trachomatis* (Table 1).

Table 1: HSV-2, HPV and *Chlamydia trachomatis* protective antigens obtained from protegen server.

Pathogen	Protegen ID	Gene Name	NCBI Protein GI	Protein Name	Molecule Role	Vaccines Involving this Gene
Herpes simplex type 2	499	US6	9629336	envelope glycoprotein D	Protective antigen	
	500	UL27 from Herpes simplex virus type 2	1869850	glycoprotein B	Protective antigen	Herpes simplex virus type 2 DNA vaccine encoding GB; Herpes simplex virus DNA vaccine ΔgB-2 encoding gB; Herpes simplex virus DNA vaccine encoding gD and gB from HSV-2 strain Curtis

	503	ICP27 from Herpes simplex virus type 2	124181	Transcriptional regulator ICP27	Protective antigen	Herpes simplex virus type 2 DNA vaccine encoding ICP27
	505	US4	9629334	envelope glycoprotein G	Protective antigen	Herpes simplex virus type 2 vaccine using phage displaying GG protein
	508	**UL44 from HSV-2**	**9629314**	**envelope glycoprotein C**	Protective antigen	
Human papillomavirus	**509**	**E7**	**9627105**	**transforming protein**	Protective antigen	Human papillomavirus E7 protein vaccine; Human papillomavirus DNA vaccine Pe7(pcDNA3-Sig/sE7/LAMP) encoding E7; Human papillomavirus DNA vaccine CRT/E7 DNA; Human papillomavirus DNA vaccine E7IR; Human papillomavirus DNA vaccine E7SH DNA; Human papillomavirus DNA vaccine hCRTE6E7L2 DNA encoding CRT linked to E6, E7 and L2; LM1-2-E7
	510	L2	9627107	minor capsid protein	Protective antigen	Human papillomavirus L2 protein vaccine; Human papillomavirus DNA vaccine hCRTE6E7L2 DNA encoding CRT linked to E6, E7 and L2
	511	E6	9627104	transforming protein	Protective antigen	Human papillomavirus DNA vaccine pNGVL4a-E6/opt; Human

						papillomavirus DNA vaccine hCRTE6E7L2 DNA encoding CRT linked to E6, E7 and L2
	512	L1	9627108	major capsid L1 protein	Protective antigen	Human papillomavirus DNA vaccine VlJns-Ll encoding L1; Human papillomavirus DNA vaccine pC16-L1 encoding L1
	1386	Calreticulin from humans	1905911	calreticulin	Protective antigen	Human papillomavirus DNA vaccine hCRTE6E7L2 DNA encoding CRT linked to E6, E7 and L2
Chlamydia trachomatis	**1682**	**ompA**	**129149**	**Major outer membrane porin, serovar E**	**Protective antigen**	*C. trachomatis* DNA vaccine pWRG7079::MOMP

Sequence collection

The complete amino acid sequences of flagellin D_1/ D_0, HPV E7 and L2, HSV-2 glycoprotein D, and Chlamydia ompA were collected from the NCBI (Table S1).

CD4+ and CD8+ T cell epitope mapping

The shared and high ranked peptide regions over multiple alleles (MHC-2Ld, MHC-2Db, MHC-2Kd, MHC-2Dd, MHC-2Kb, and MHC-2Kk) were selected from each server for both Glyco-D and OmpA proteins (Table S2). Using the IEBD server, the obtained epitopes were chosen as final CD8+ epitopes. Additionally, Glyco-D and OmpA were employed for determining the high-affinity CD4+ epitopes binding to MHC-2IAs, MHC-2IAb, MHC-2IEd, and MHC-2IAd alleles. The proteins were analyzed with the IEDB, MHCpred, and RANKPEP servers (Tables S3). MHCpred uses QSAR stimulation approach for the calculation of the half maximal (50%)

inhibitory concentration (IC_{50}) for each epitope. Generally, IC_{50} values range between 0.01 and 5000 nM demonstrate high binding affinity of peptide to MHC-II. Peptides with lower IC_{50} values are good binders and vice versa, and IC_{50} values above 5000 indicate that peptides will not bind to MHC-II molecules.

Prediction of TCR-peptide/peptide-MHC interfaces and CTL epitopes

Determination of T cell epitopes was performed by PAComplex and CTLpred for the Glyco-D and OmpA proteins (Tables S4). Joint Z-value is used in PAComplex for defining CTL epitopes, and higher joint Z-values signify more affinity for both MHC and TCR molecules. This server uses only four mouse alleles for the identification of MHC-2Ld and MHC-2Kb epitopes. Prior to epitope selection, the signal peptide sequences were excluded from the selected protective antigens. The conserved CD4+ and CD8+ T cell epitopes of HSV-2, CT ompA, glycoprotein C (gC) from the current study, and two regions of HPV containing the best epitopes of E7 and L2 protein, based on the previous research, were selected (Table 2).

Table 2: The common CD4+ and CD8+ T cell epitopes from HSV-2 gC, HPV E7, and *Ct* ompA

Protective Antigens	Immunodominant epitopes	Start aa	End aa
HSv-2 gC	RYGSRVQIRCRFPNSTRTEFRLQIWRYATATDAEIGTAPSLEEVMVNV	77	124
	RTDRPSAYGTWVRVR	190	204
HPV E7, L2	GTGGRTGYIPLGTRPPTATDTLAPV	65	89
	EPDRAHYNIVTFCCKCDSTL	46	65
Ct ompA	GTKDASIDYHEWQASLALSYRLNMFTPYIG	238	267

Construction of the epitope vaccine

The schematic representation of the vaccine construct consists of flagellin D_1/D_0 immunodominant, selected epitopes from HPV E7 and L2, HSV-2 gC, *Ct* ompA, P30 as helper adjuvant, and RS09 (TLR4 agonist), which are joined together by proper linkers (Fig. 1).

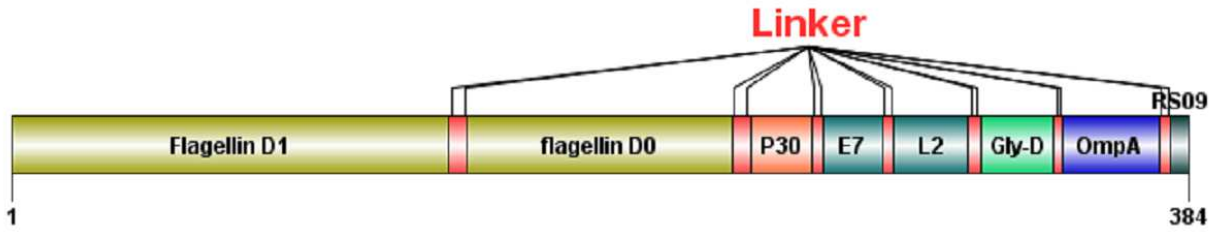

Fig. 1: Schematic diagram of vaccine sequence consists of conserved amino acid residues of flagellin (1–143,149-235), universal helper P30 (241-261) and protective antigens consist E7 and L2 domain of HPV (265-284, 288-312), Glyco-D (316-340), OmpA (343-374) and RS09 (378-384) which fused together by proper linkers.

Vaccine features

Solubility, antigenicity, allergenicity, and physicochemical parameters evaluation

ANTIGENpro and VaxiJen estimated the probability of the whole vaccine antigenicity 0.88% and 0.6031 (threshold is 0.4), respectively. Prediction of the allergenicity by AllergenFP v.1.0 server confirmed the vaccine is a non-allergen.

The Solpro server predicted the protein was soluble with the probability of 0.646909, and according to the result of Protein-sol, the scaled solubility value (QuerySol) was predicted to be 0.425. The population average is 0.45 and scaled solubility value is near to this average and is predicted to soluble. The instability index (II) was calculated to be 34.84, which classifies the protein as stable. The number of amino acids was 384, and the molecular weight (Mw) and theoretical isoelectric point value (pI) of the protein were computed 41.07 kDa and 8.92, respectively. Half-life was estimated to be 30 hours in mammalian reticulocytes, >20 hours in yeasts, and >10 h in *E. coli*. Aliphatic index and GRAVY were estimated 83.93 and -0.381, respectively.

Tertiary structure refinement and homology modeling

I-TASSER program was used to build the 3D model of the fusion protein vaccine. Five 3D models of our protein vaccine were obtained from the server (with c-scores: -0.14, -0.63, -1.93, -3.49, and -3.84). C-score is a confidence score for determining the quality of the models predicted by I-TASSER. The potential errors and quality in an initial 3D model were established by Ramachandran, verify 3D, Prosa-web and ERRAT. The verify 3D model results indicated that the crude 3D model requires refinement processes. Thus, according to result of validation by Ramachandran, ERRAT and Verify 3D, the model 2 of the I-TASSER model was employed for this purpose (Table S8). The GalaxyRefine web server was employed for the refinement process. According to result of validation by Ramachandran plot, Verify 3D, ERRAT and Prosa-web the

most promising model was employed for whole refinement and energy minimization. The best-quality refined model (model 3) was chosen after all the refinement steps (Table 3). TMB-refine and loop-refine were performed and the results are shown in tables S9 and S10.

Table 3: Validation refined models 0f model-2 / I-TASSER by Galaxyrefine server

Server	Prosaweb (z score)	Ramachandran Plot	ERRAT (Overall Quality Factor)	Verify 3D
Model 1	-4.29	residues in favoured region: 92.1 residues in allowed region: 6.5 residues in outlier region: 1.3	85.50	46.35
Model 2	-4.62	residues in favoured region: 92.1 residues in allowed region: 6.3 residues in outlier region: 1.6	81.36	52.34
Model 3	**-4.41**	**residues in favoured region: 92.4** **residues in allowed region: 5.8** **residues in outlier region: 1.8**	**86.97**	**48.44**
Model 4	-4.39	residues in favoured region: 92.1 residues in allowed region: 6.0 residues in outlier region: 1.8	85.54	48.70
Model 5	-4.52	residues in favoured region: 92.1 residues in allowed region: 6.0 residues in outlier region: 1.8	84.63	55.99

Defining B-cell epitopes

Linear B-cell epitopes were identified by BCEPRED and ABCPred. In ABCPRED, 20 mer length and 75% specificity index of B-cell epitopes were used as the default parameters, and 20 mer length of B-cell epitopes and cut-off of 2.38 were set for BCEPRED. The 28 epitopes identified by ABCPRED are presented in Table S5, and the 10 epitopes determined by BCEPRED are shown in Table S6. In addition, the 3D model of the protein was used to determine discontinuous B-cell epitopes by the Disco tope 2.0 server (Table S7).

Molecular dynamic simulation of the multi-epitope vaccine

The MD simulation program with the Amber ff99SB*-ILDN force-field was applied to the vaccine construct obtained in the previous step. Energy minimization was performed for 50000 steps, and upon reaching 3568 steps, the steepest descents converged, and the force attained was <1000 KJ/mol. The system was equilibrated so that the kinetic and potential energy changes showed stable behavior during simulation. Density, pressure, and temperature became stabilized at around 993 kg/m^3, 300 K (Fig. 2A), and 1 bar (Fig. 2B), respectively. Trajectory was analyzed after 50 ns simulation time. Then, the backbone RMSD of the multi-epitope vaccine was graphed. The RMSD plot revealed that RMSD levels increased to ~1.9 nm and stay between 1.9 nm and 1.4 nm throughout the simulation time, showing that the structure is highly stable (Fig. 2C). On the other hand, RMSF signifies regions with high flexibility (Fig. 2D). The residues 37-53, 93-115, 146-153, 182-188, 196-230, 234-246, 251-255, 261-267, and 297-308 of the designed vaccine had high fluctuations. It is worth mentioning that all the residues of the GPGPG linker, and its neighbors (146-153) showed significant changes in their fluctuations. Furthermore, all the regions, including KK, and GGS linker had mild fluctuations, except for the C-terminus GGS linker (375-377). Regions including the EAAAK linker had an insignificant peak in the RMSF plot.

Additionally, computing solvent accessible surface area (SASA) can be used to determine alterations in the accessibility of protein to solvent. The average SASA value of the vaccine molecule showed the stability and compactness of its structure (Fig. 2E). Radius of gyration, which is a key parameter indicating compactness of structure, also exhibited that the vaccine molecule has gained a stable and compact form during MD simulation (Fig. 2F).

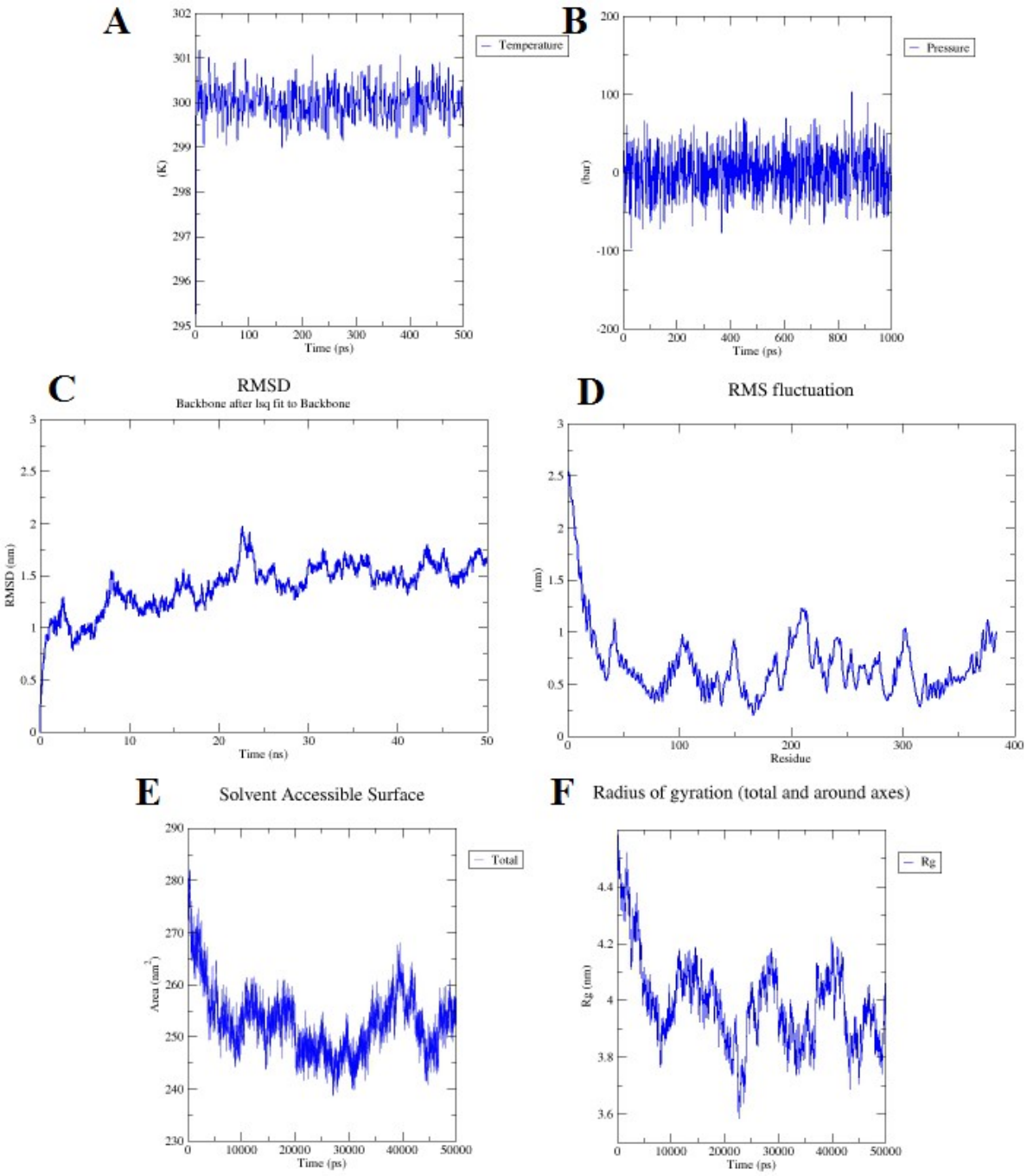

Fig. 2: Molecular dynamics simulation of the vaccine candidate (**A**) Temperature; system temperature reached to 300 K and stay constant around 300 K (500 ps) (**B**) Pressure; Ligand pressure plot at equilibration phase of 1000 ps. (**C**) RMSD; RMSD levels represents the stability of vaccine construct (**D**) RMSF; RMSF-Root Mean Square Fluctuation plot, peaks show the regions with high flexibility. (E) solvent accessible surface area (SASA) plot, vaccine construct

is compact and stable during the simulation time. (**F**) Rg plot; vaccine construct is stable in its compact form during the simulation time

Molecular docking of the vaccine construct with Toll-like receptor 5

Vaccine construct and TLR5 docking was analyzed by the ClusPro server. The output of this server is composed of multiple models in four force-based categories, and analysis of the results is based on free binding energies and cluster size. The docking analysis signified suitable interaction between TLR5, and the vaccine construct. In Fig. 3A, TLR5 is exhibited in purple, and the vaccine construct is displayed in red. Furthermore, the PDB sum online database was used to determine the interactions pattern between the docked complex. This interactions pattern contained hydrogen bond and nonbonded interactions between docked proteins complex. Our vaccine construct developed nine hydrogen bond interactions (Chain A [TLR5]-B [vaccine construct]; 431-84, 432-94, 480-119, 480-119, 480-84, 509-125, 509-125, 529-125, and 532-115) with TLR5 (Fig. 3B). The interface area (Å2) of the TLR5 and vaccine was found to be 1045. Collectively, 20 interface residues of the vaccine construct were associated with 18 TLR5 residues (Fig. 3C). Based on structural analysis, the majority of the hydrogen bonds between the vaccine and TLR5 were within the range of 2 Å to 3 Å (Fig. 4), demonstrating robust interactions [38].

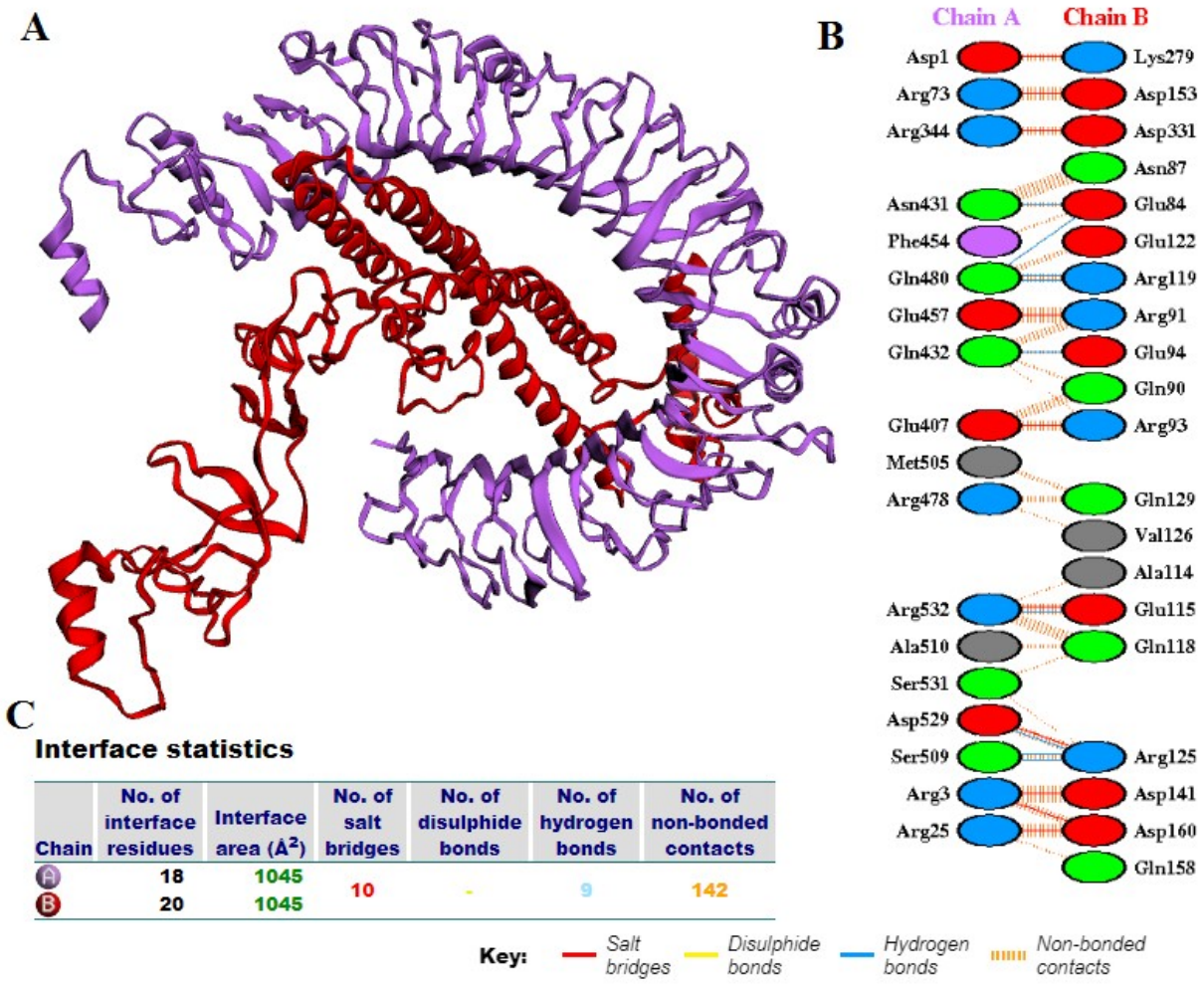

Fig. 3: Vaccine construct-TLR5 docked complex: A) Figure obtained after molecular docking between Vaccine and TLR-3, red color showed the vaccine construct while purple color is representing TLR5. B) The interacting residues illustration between docked vaccine (chain B) and TLR5 (chain A) complex. C) Interface statics of vaccine construct-TLR5 docked complex.

Hydrogen bonds

	<----- A T O M 1 ----->					<----- A T O M 2 ----->				
	Atom no.	Atom name	Res name	Res no.	Chain		Atom no.	Atom name	Res name	Res no.	Chain	Distance
1.	4222	ND2	ASN	431	A	<-->	770	OE1	GLU	84	B	2.91
2.	4234	NE2	GLN	432	A	<-->	885	OE1	GLU	94	B	2.91
3.	4699	OE1	GLN	480	A	<-->	1116	NH1	ARG	119	B	3.08
4.	4699	OE1	GLN	480	A	<-->	1119	NH2	ARG	119	B	2.64
5.	4700	NE2	GLN	480	A	<-->	771	OE2	GLU	84	B	3.26
6.	4979	OG	SER	509	A	<-->	1181	NH1	ARG	125	B	3.25
7.	4979	OG	SER	509	A	<-->	1184	NH2	ARG	125	B	2.76
8.	5151	OD1	ASP	529	A	<-->	1181	NH1	ARG	125	B	2.79
9.	5178	NE	ARG	532	A	<-->	1073	OE1	GLU	115	B	2.90

Fig. 4: Coordinates of hydrogen bonds between vaccine construct and receptor TLR5.

Discussion

According to WHO records, the prevalence of HSV type 1, HSV type 2, and genital HSV type 1 infections in 2016 has been reported 3583, 491, and 122-192 million cases, respectively. Different patterns of this infection are observed by sex, geographical region, and age, but it is more common in the female population and African nations [39].

Two target populations have been considered for HSV-2 vaccine development, which include those already infected with HSV-2 and those who are not. For the first group, therapeutic vaccines and for the second prophylactic vaccines are designed.

Several Phase III clinical trials have been performed to evaluate the efficacy of prophylactic subunit vaccines based on glycoprotein D2. However, none of them prevented HSV-2 acquisition despite eliciting strong neutralizing antibody responses. Recently, HerpeVac project showed that adjuvanted gD2 was unable to prevent symptomatic genital herpes [40]. It has been proposed that vaccine efficacy against HSV-1 can be enhanced by increasing antibody titers to gD2 [41]. Attempts to develop prophylactic HSV vaccines declined following these studies, but the development of a therapeutic HSV-2 vaccine has attracted considerable interest in recent years. Current vaccine candidates target HSV-2, but vaccines targeting both HSV types could be developed due to the identification of conserved epitopes in both HSV-1 and HSV-2 [42].

In the context of chlamydia, over 130 million new cases of C*hlamydia trachomatis* infection are reported every year [1]. This type of infection can influence female reproductive health and lead to problems such as repeated and spontaneous abortions, ectopic pregnancies, and stillbirths [43, 44]. HPV and *C. trachomatis* coinfection elevate the risk of cervical cancer [45]. Li et al.

reported the prevalence of HPV, *C. trachomatis*, and *C. trachomatis*/HPV coinfection to be 36.0, 14.3, and 4.8%, respectively, in the northern part of Inner Mongolia, China [46].

The majority of sexually active women and men experience HPV infection as least once in their lifetime. There are different types of HPVs, but two of them (16 and 18) cause the majority of pre-cancerous cervical lesions and cervical cancers (70%). Cervical cancer is very common in the female population living in the developing countries with an estimated incidence rate of 570,000 cases [47] in 2018 (84% of the new cases worldwide). In that year, cervical cancer caused about 311,000 deaths in the female population, only 15% of which occurred in high-income countries.

Since 2012, studies have shown that sexually transmitted infections are persistently endemic worldwide [48]. Two decades age, STI vaccines against HPV and hepatitis B seemed impossible [49], while now a new generation of STI vaccines is underway, and several vaccine candidates in early clinical trials have shown promising results. Many STI vaccine candidates are being examined in phase I trial, and the new chlamydia vaccine candidate is the case in point. New syphilis outbreaks and antimicrobial resistance for gonorrhea are among the emerging challenges making the development these vaccines an urgency which can be met most quickly by collaboration and support [49-51].

The evolution of pathogenic microbes has posed a serious challenge to vaccine development. During the evolutionary process, pathogenic microbes produce a variety of antigens and create a barrier to vaccine development. This problem has tackled by incorporating different antigenic variants into a single vaccine. Recently, different combinatorial forms of subunit vaccine against two or three types of bacteria have been designed; for instance, in 2013 Santana et al. developed the multi-epitope DNA vaccine harboring HSV, HPV and HIV antigens that could elicit efficient humoral and cellular responses [11]. Moreover, in 2006 Macmillan et al. Designed a vaccine against Chlamydia and HSV-2 [52].

Various protective antigens are used for designing a vaccine against STI. Investigations have shown the important role of HPV 16 E6, E7, and L1 and L2N in cervical cancer and HPV infection [53-59]. Envelope glycoproteins (i.e., gD, gB, gH, gL, and gC) are well-known protective antigens for HSV-2 vaccine candidates. A previous study revealed that efficacy against HSV1 is significantly correlated with gD concentration [41]. Karunakaran et al. established the importance of Chlamydia outer membrane proteins for vaccine development

against *Chlamidia trachomatis* [60], and Vasilevsky et al. approved that polymorphic membrane proteins (Pomps) could be used as a vaccine candidate against *C. trachomatis* because of its main role in the pathogenicity of chlamydial infection [61].

The problem of poor immunogenicity of subunit vaccines can be overcome by the use of adjuvants, in this context, TLR agonists are among the main adjuvants that stimulate the immune system directly and indirectly as well as can facilitate entry of protein to cell. Flagellin is a bacterial protein essential for bacterial motility. This protein can be turned into a filament unit by the polymerization process. As flagellated bacteria invade the host, TLR5 identifies flagellin as a pathogen invasion signal and stimulates the innate immune response. Flagellin is used as a vaccine delivery protein or a vaccine adjuvant as it triggers the first line of immunity against flagellated pathogenic bacteria [62]. In this regard, we chose D0/D1 domains as important regions of flagellin in our final construct; these domains are critical for motility-related functions and polymerization of the flagellum [63].

The RS09 was introduced as TLR4 agonistic peptide motif by Masaki et al., who investigated whether these antigens could induce antigen-specific CTL response and the maturation of antigen presenting cells (APCs) without using any conventional adjuvants. They noted the production of TNF-α and the expression of CD40 in APC are induced by this motif. This finding corroborates the maturation of APC *in-vitro* [64]. Smahel et al. showed the essential role of cellular localization of modified antigens for the impact of the p30 epitope, which is linked with T_h2 response activation. P30 was fused with E7GGG, which could contribute to its proper processing and presentation with MHC class II molecules [65].

In this research, first the important protective antigens of HSV-2, HPV, and *Chlamydia trachomatis* were retrieved from the VIOLIN server. E7 and L2 from HPV, glycoprotein D from HSV-2, and ompA from Ct were selected as the final antigenic proteins. The immunodominant HTL and CTL epitopes were selected from the aforementioned antigens by using the appropriate data bank. Epitope selection is based on conservancy and immunogenicity of the regions, and according to these parameters the highest rank immunogenic and conserved epitopes were chosen. The P30 (tetanus toxoid epitope), as a universal T-helper, was fused to E7, which can bind to multiple human and mouse MHC-II alleles and stimulate T-helper cell responses.

Moreover, flagellin D_1/D_0 as TLR5 agonist and RS09 as a TLR4 agonist were added to N and C-terminals of the peptide vaccine, respectively. In the next stage, all parts of the vaccine were

fused together by appropriate linkers to enhance vaccine efficiency. Evaluation of antigenicity and allergenicity showed that this chimeric protein was a suitable antigen and non-allergen. Physicochemical properties assessment demonstrated this protein was soluble and stable with proper hydrophobicity and hydrophilicity.

The RMSD plot showed that the protein vaccine structure was very stable (Fig. 2C). Furthermore, as shown in Fig. 2E, the calculated average SASA value of the vaccine molecule indicated the compactness and stability of its structure. Radius of gyration also showed that vaccine molecule had become structurally stable and compact during MD simulation (Fig. 2F). Evaluation of molecular docking between protein and TLR5 demonstrated that most of the hydrogen bonds between the vaccine, and TLR5 have strong interactions. Various regions of final construct were identified as linear and discontinuous B-cell epitopes that could potentially stimulate the immune system via raising antibody secretion by B-cells.

5. Conclusion

In the present study, a multi-epitope vaccine was designed against *C. trachomatis*, HPV, and HSV2, which are the main agents of STDs. This vaccine can cause clearance of these pathogens and produce immunity via increasing the secretion of antibodies through stimulating B-cells and a mixed Th1/Th2 cells response. This study was carried out based on reverse vaccinology by adopting the immunoinformatics methods. However, further in vitro and in vivo immunological studies are warranted to corroborate the efficacy of this epitope vaccine.

Acknowledgements

This work was supported by Shiraz University of Medical Science grant, Shiraz, Iran.

Authors' contributions

Conceptualization: HD, Methodology: MZ, and SE, Drafting the manuscript: HD, NN, and YG, MS and RJ, Revising the manuscript critically for important intellectual content: NN and YG, Supervision: NN, and YG. All authors have read and approved the manuscript.

Funding

This study was funded by Shiraz University of Medical Science grant.

Availability of data and materials

Input data for the analyses are available from the corresponding author on request.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interest

References:

1. Newman L, Rowley J, Vander Hoorn S, Wijesooriya NS, Unemo M, Low N, Stevens G, Gottlieb S, Kiarie J, Temmerman M: **Global Estimates of the Prevalence and Incidence of Four Curable Sexually Transmitted Infections in 2012 Based on Systematic Review and Global Reporting**. *PLoS One* 2015, **10**(12):e0143304.
2. James C, Harfouche M, Welton NJ, Turner KM, Abu-Raddad LJ, Gottlieb SL, Looker KJ: **Herpes simplex virus: global infection prevalence and incidence estimates, 2016**. *Bulletin of the World Health Organization* 2020, **98**(5):315-329.
3. Organization WH: **Report on global sexually transmitted infection surveillance 2018**. 2018.
4. Brunham RC, Gottlieb SL, Paavonen J: **Pelvic inflammatory disease**. *The New England journal of medicine* 2015, **372**(21):2039-2048.

5. Ohshige K, Morio S, Mizushima S, Kitamura K, Tajima K, Suyama A, Usuku S, Tia P, Hor L, Heng S *et al*: **Behavioural and serological human immunodeficiency virus risk factors among female commercial sex workers in Cambodia**. *International Journal of Epidemiology* 2000, **29**(2):344-354.

6. Trimble CL, Frazer IH: **Development of therapeutic HPV vaccines**. *The Lancet Oncology* 2009, **10**(10):975-980.

7. Johnston C, Koelle DM, Wald A: **Current status and prospects for development of an HSV vaccine**. *Vaccine* 2014, **32**(14):1553-1560.

8. Truong NR, Smith JB, Sandgren KJ, Cunningham AL: **Mechanisms of Immune Control of Mucosal HSV Infection: A Guide to Rational Vaccine Design**. *Frontiers in Immunology* 2019, **10**(373).

9. Haghshenas MR, Mousavi T, Kheradmand M, Afshari M, Moosazadeh M: **Efficacy of Human Papillomavirus L1 Protein Vaccines (Cervarix and Gardasil) in Reducing the Risk of Cervical Intraepithelial Neoplasia: A Meta-analysis**. *Int J Prev Med* 2017, **8**:44-44.

10. Phillips S, Quigley BL, Timms P: **Seventy Years of Chlamydia Vaccine Research – Limitations of the Past and Directions for the Future**. *Frontiers in Microbiology* 2019, **10**(70).

11. Santana VC, Diniz MO, Cariri FA, Ventura AM, Cunha-Neto E, Almeida RR, Campos MA, Lima GK, Ferreira LC: **Bicistronic DNA vaccines simultaneously encoding HIV, HSV and HPV antigens promote CD8(+) T cell responses and protective immunity**. *PLoS One* 2013, **8**(8):e71322.

12. Karanam B, Jagu S, Huh WK, Roden RBS: **Developing vaccines against minor capsid antigen L2 to prevent papillomavirus infection**. *Immunol Cell Biol* 2009, **87**(4):287-299.

13. Cheng MA, Farmer E, Huang C, Lin J, Hung C-F, Wu TC: **Therapeutic DNA Vaccines for Human Papillomavirus and Associated Diseases**. *Hum Gene Ther* 2018, **29**(9):971-996.

14. El-Zayat SR, Sibaii H, Mannaa FA: **Toll-like receptors activation, signaling, and targeting: an overview**. *Bulletin of the National Research Centre* 2019, **43**(1):187.

15. Fazeli A, Bruce C, Anumba DO: **Characterization of Toll-like receptors in the female reproductive tract in humans**. *Human reproduction (Oxford, England)* 2005, **20**(5):1372-1378.

16. Vita R, Overton JA, Greenbaum JA, Ponomarenko J, Clark JD, Cantrell JR, Wheeler DK, Gabbard JL, Hix D, Sette A *et al*: **The immune epitope database (IEDB) 3.0**. *Nucleic acids research* 2015, **43**(Database issue):D405-D412.

17. Guan P, Doytchinova IA, Zygouri C, Flower DR: **MHCPred: A server for quantitative prediction of peptide-MHC binding**. *Nucleic Acids Res* 2003, **31**(13):3621-3624.

18. Reche PA, Glutting JP, Reinherz EL: **Prediction of MHC class I binding peptides using profile motifs**. *Human immunology* 2002, **63**(9):701-709.

19. Liu IH, Lo YS, Yang JM: **PAComplex: a web server to infer peptide antigen families and binding models from TCR-pMHC complexes**. *Nucleic Acids Res* 2011, **39**(Web Server issue):W254-260.

20. Saha S, Raghava GP: **Prediction of continuous B-cell epitopes in an antigen using recurrent neural network**. *Proteins* 2006, **65**(1):40-48.

21. Saha S, Raghava GPS: **BcePred: Prediction of Continuous B-Cell Epitopes in Antigenic Sequences Using Physico-chemical Properties**. In: *Artificial Immune Systems: 2004// 2004; Berlin, Heidelberg*. Springer Berlin Heidelberg: 197-204.

22. Kringelum JV, Lundegaard C, Lund O, Nielsen M: **Reliable B cell epitope predictions: impacts of method development and improved benchmarking**. *PLoS computational biology* 2012, **8**(12):e1002829.

23. Magnan CN, Zeller M, Kayala MA, Vigil A, Randall A, Felgner PL, Baldi P: **High-throughput prediction of protein antigenicity using protein microarray data**. *Bioinformatics* 2010, **26**(23):2936-2943.

24. Doytchinova IA, Flower DR: **Bioinformatic approach for identifying parasite and fungal candidate subunit vaccines**. *Open Vaccine J* 2008, **1**(1):4.

25. Saha S, Raghava GP: **AlgPred: prediction of allergenic proteins and mapping of IgE epitopes**. *Nucleic Acids Res* 2006, **34**(Web Server issue):W202-209.

26. Dimitrov I, Bangov I, Flower DR, Doytchinova I: **AllerTOP v.2--a server for in silico prediction of allergens**. *Journal of molecular modeling* 2014, **20**(6):2278.

27. Dorosti H, Eslami M, Nezafat N, Fadaei F, Ghasemi Y: **Designing self-assembled peptide nanovaccine against Streptococcus pneumoniae: An in silico strategy**. *Molecular and cellular probes* 2019, **48**:101446-101446.

28. Gasteiger E, Hoogland C, Gattiker A, Wilkins MR, Appel RD, Bairoch A: **Protein identification and analysis tools on the ExPASy server**. In: *The proteomics protocols handbook*. Springer; 2005: 571-607.

29. Hebditch M, Carballo-Amador MA, Charonis S, Curtis R, Warwicker J: **Protein–Sol: a web tool for predicting protein solubility from sequence**. *Bioinformatics* 2017, **33**(19):3098-3100.

30. Vakili B, Eslami M, Hatam GR, Zare B, Erfani N, Nezafat N, Ghasemi Y: **Immunoinformatics-aided design of a potential multi-epitope peptide vaccine against Leishmania infantum**. *International Journal of Biological Macromolecules* 2018, **120**:1127-1139.

31. Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y: **The I-TASSER Suite: protein structure and function prediction**. *Nature methods* 2015, **12**(1):7.

32. Ko J, Park H, Heo L, Seok C: **GalaxyWEB server for protein structure prediction and refinement**. *Nucleic acids research* 2012, **40**(Web Server issue):W294-W297.

33. Heo L, Park H, Seok C: **GalaxyRefine: protein structure refinement driven by side-chain repacking**. *Nucleic acids research* 2013, **41**(W1):W384-W388.

34. Heo L, Lee H, Seok C: **GalaxyRefineComplex: Refinement of protein-protein complex model structures driven by interface repacking**. *Sci Rep* 2016, **6**:32153-32153.

35. Wiederstein M, Sippl MJ: **ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins**. *Nucleic acids research* 2007, **35**(suppl_2):W407-W410.

36. Kozakov D, Hall DR, Xia B, Porter KA, Padhorny D, Yueh C, Beglov D, Vajda S: **The ClusPro web server for protein–protein docking**. *Nature protocols* 2017, **12**(2):255.

37. Laskowski RA, Jabłońska J, Pravda L, Vařeková RS, Thornton JM: **PDBsum: Structural summaries of PDB entries**. *Protein Sci* 2018, **27**(1):129-134.

38. Dannenberg J: **An Introduction to Hydrogen Bonding By George A. Jeffrey (University of Pittsburgh). Oxford University Press: New York and Oxford. 1997. ix+ 303 pp. $60.00. ISBN 0-19-509549-9**. In.: ACS Publications; 1998.

39. James C, Harfouche M, Welton NJ, Turner KM, Abu-Raddad LJ, Gottlieb SL, Looker KJ: **Herpes simplex virus: global infection prevalence and incidence estimates, 2016**. *Bull World Health Organ* 2020, **98**:315-329.

40. Zhu J, Peng T, Johnston C, Phasouk K, Kask AS, Klock A, Jin L, Diem K, Koelle DM, Wald A *et al*: **Immune surveillance by CD8αα+ skin-resident T cells in human herpes virus infection**. *Nature* 2013, **497**(7450):494-497.

41. Belshe RB, Heineman TC, Bernstein DI, Bellamy AR, Ewell M, van der Most R, Deal CD: **Correlate of immune protection against HSV-1 genital disease in vaccinated women**. *The Journal of infectious diseases* 2014, **209**(6):828-836.

42. Jing L, Laing KJ, Dong L, Russell RM, Barlow RS, Haas JG, Ramchandani MS, Johnston C, Buus S, Redwood AJ: **Extensive CD4 and CD8 T cell cross-reactivity between alphaherpesviruses**. *The Journal of Immunology* 2016, **196**(5):2205-2218.

43. Tiller CM: **Chlamydia during pregnancy: implications and impact on perinatal and neonatal outcomes**. *Journal of Obstetric, Gynecologic, & Neonatal Nursing* 2002, **31**(1):93-98.

44. Baud D, Greub G: **Intracellular bacteria and adverse pregnancy outcomes**. *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases* 2011, **17**(9):1312-1322.

45. Naldini G, Grisci C, Chiavarini M, Fabiani R: **Association between human papillomavirus and chlamydia trachomatis infection risk in women: a systematic review and meta-analysis**. *International journal of public health* 2019:1-13.

46. Ji Y, Ma X-x, Li Z, Peppelenbosch MP, Ma Z, Pan Q: **The burden of human papillomavirus and Chlamydia trachomatis coinfection in women: a large cohort study in inner Mongolia, China**. *The Journal of infectious diseases* 2019, **219**(2):206-214.

47. Ferlay J, Colombet M, Soerjomataram I, Mathers C, Parkin DM, Piñeros M, Znaor A, Bray F: **Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods**. *International journal of cancer* 2019, **144**(8):1941-1953.

48. Rowley J, Vander Hoorn S, Korenromp E, Low N, Unemo M, Abu-Raddad LJ, Chico RM, Smolak A, Newman L, Gottlieb S *et al*: **Chlamydia, gonorrhoea, trichomoniasis and syphilis: global prevalence and incidence estimates, 2016**. *Bulletin of the World Health Organization* 2019, **97**(8):548-562P.

49. Gottlieb SL, Johnston C: **Future prospects for new vaccines against sexually transmitted infections**. *Curr Opin Infect Dis* 2017, **30**(1):77-86.

50. Broutet N, Fruth U, Deal C, Gottlieb SL, Rees H: **Vaccines against sexually transmitted infections: the way forward**. *Vaccine* 2014, **32**(14):1630-1637.

51. Gottlieb SL, Deal CD, Giersing B, Rees H, Bolan G, Johnston C, Timms P, Gray-Owen SD, Jerse AE, Cameron CE *et al*: **The global roadmap for advancing development of vaccines against sexually transmitted infections: Update and next steps**. *Vaccine* 2016, **34**(26):2939-2947.

52. Macmillan L, Ifere GO, He Q, Igietseme JU, Kellar KL, Okenu DM, Eko FO: **A recombinant multivalent combination vaccine protects against Chlamydia and genital herpes**. *FEMS immunology and medical microbiology* 2007, **49**(1):46-55.

53. Nindl I, Rindfleisch K, Lotz B, Schneider A, Dürst M: **Uniform distribution of HPV 16 E6 and E7 variants in patients with normal histology, cervical intra-epithelial neoplasia and cervical cancer**. *International journal of cancer* 1999, **82**(2):203-207.

54. Sathish N, Abraham P, Peedicayil A, Sridharan G, Chandy G: **HPV 16 E6 sequence variations in Indian patients with cervical neoplasia**. *Cancer letters* 2005, **229**(1):93-99.

55. Tornesello ML, Duraturo ML, Salatiello I, Buonaguro L, Losito S, Botti G, Stellato G, Greggi S, Piccoli R, Pilotti S: **Analysis of human papillomavirus type-16 variants in Italian women with cervical intraepithelial neoplasia and cervical cancer**. *Journal of medical virology* 2004, **74**(1):117-126.

56. Wu Y, Chen Y, Li L, Yu G, He Y, Zhang Y: **Analysis of mutations in the E6/E7 oncogenes and L1 gene of human papillomavirus 16 cervical cancer isolates from China**. *Journal of general virology* 2006, **87**(5):1181-1188.

57. Yamada T, Manos MM, Peto J, Greer CE, Munoz N, Bosch FX, Wheeler CM: **Human papillomavirus type 16 sequence variation in cervical cancers: a worldwide perspective**. *Journal of virology* 1997, **71**(3):2463-2472.

58. Auvinen E, Kujari H, Arstila P, Hukkanen V: **Expression of the L2 and E7 genes of the human papillomavirus type 16 in female genital dysplasias**. *The American journal of pathology* 1992, **141**(5):1217-1224.

59. Yan H, Foo SS, Chen W, Yoo JS, Shin WJ, Wu C, Jung JU: **Efficient Inhibition of Human Papillomavirus Infection by L2 Minor Capsid-Derived Lipopeptide**. *mBio* 2019, **10**(4).

60. Karunakaran KP, Yu H, Jiang X, Chan Q, Moon KM, Foster LJ, Brunham RC: **Outer membrane proteins preferentially load MHC class II peptides: implications for a Chlamydia trachomatis T cell vaccine**. *Vaccine* 2015, **33**(18):2159-2166.

61. Vasilevsky S, Stojanov M, Greub G, Baud D: **Chlamydial polymorphic membrane proteins: regulation, function and potential vaccine candidates**. *Virulence* 2016, **7**(1):11-22.

62. Mizel SB, Bates JT: **Flagellin as an adjuvant: cellular mechanisms and potential**. *The Journal of Immunology* 2010, **185**(10):5677-5682.

63. Schijns VEJC: **Chapter 1 - Vaccine Adjuvants' Mode of Action: Unraveling ''the Immunologist's Dirty Little Secret"**. In: *Immunopotentiators in Modern Vaccines (Second Edition)*. Edited by Schijns VEJC, O'Hagan DT: Academic Press; 2017: 1-22.

64. Ito M, Hayashi K, Minamisawa T, Homma S, Koido S, Shiba K: **Encryption of agonistic motifs for TLR4 into artificial antigens augmented the maturation of antigen-presenting cells**. *PLOS ONE* 2017, **12**(11):e0188934.

65. Wolkers MC, Toebes M, Okabe M, Haanen JB, Schumacher TN: **Optimizing the efficacy of epitope-directed DNA vaccination**. *The Journal of Immunology* 2002, **168**(10):4998-5004.

Figures

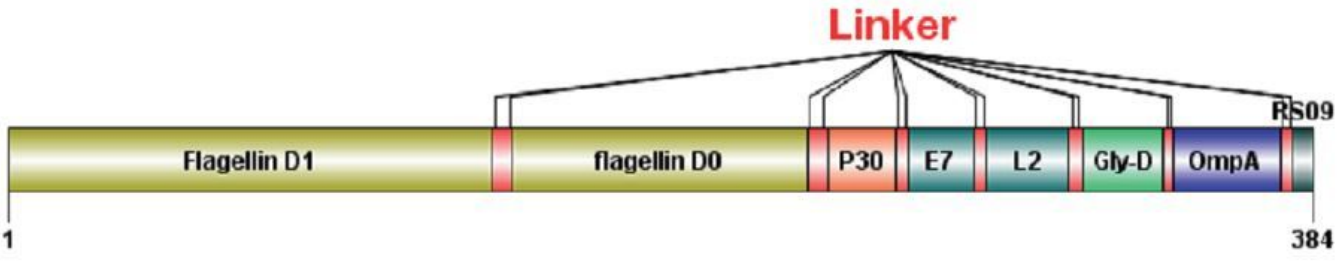

Figure 1

Schematic diagram of vaccine sequence consists of conserved amino acid residues of flagellin (1–143,149-235), universal helper P30 (241-261) and protective antigens consist E7 and L2 domain of HPV (265-284, 288-312), Glyco-D (316-340), OmpA (343-374) and RS09 (378-384) which fused together by proper linkers.

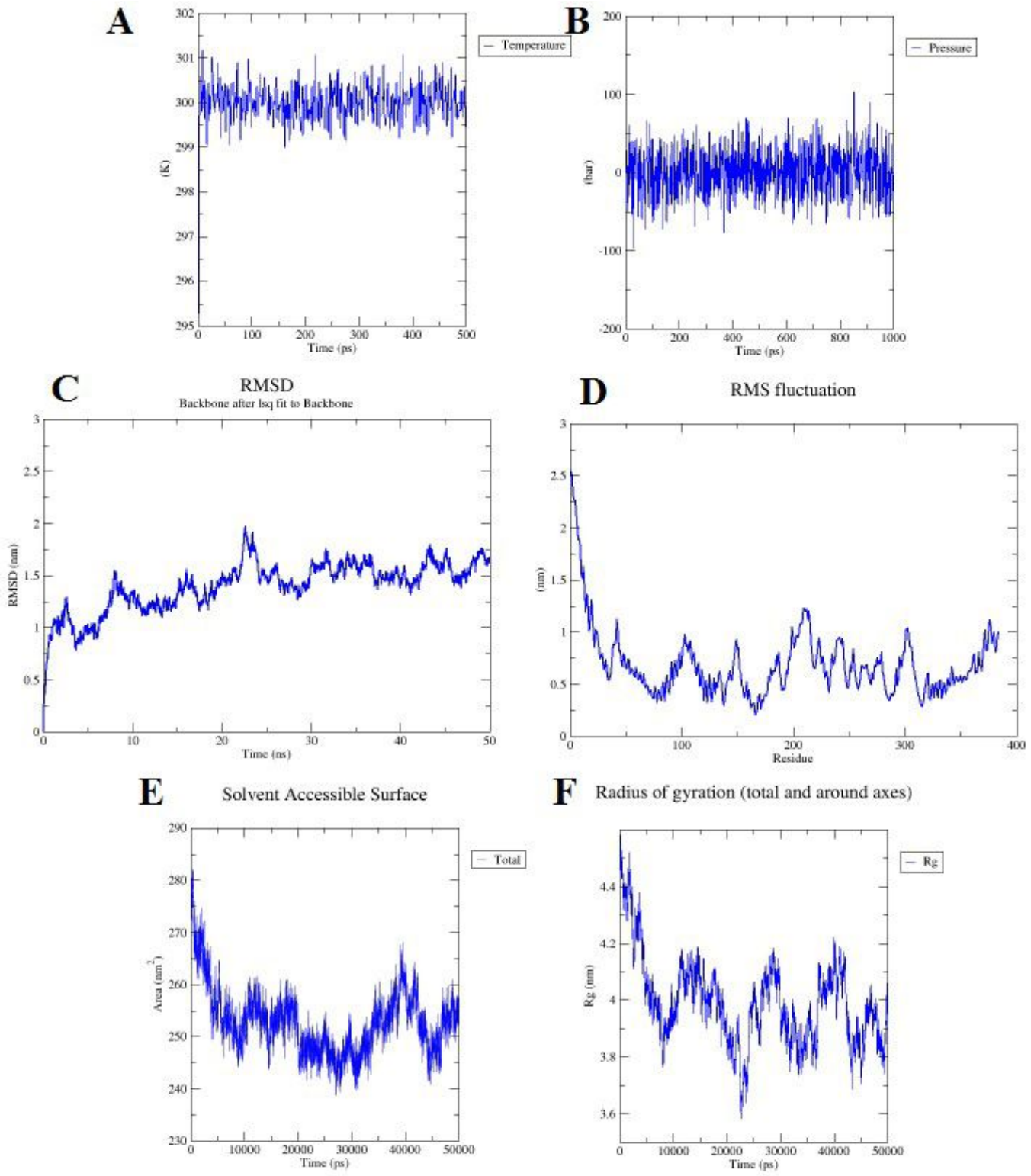

Figure 2

Molecular dynamics simulation of the vaccine candidate (A) Temperature; system temperature reached to 300 K and stay constant around 300 K (500 ps) (B) Pressure; Ligand pressure plot at equilibration phase of 1000 ps. (C) RMSD; RMSD levels represents the stability of vaccine construct (D) RMSF; RMSF-Root Mean Square Fluctuation plot, peaks show the regions with high flexibility. (E) solvent accessible surface

area (SASA) plot, vaccine construct is compact and stable during the simulation time. (F) Rg plot; vaccine construct is stable in its compact form during the simulation time

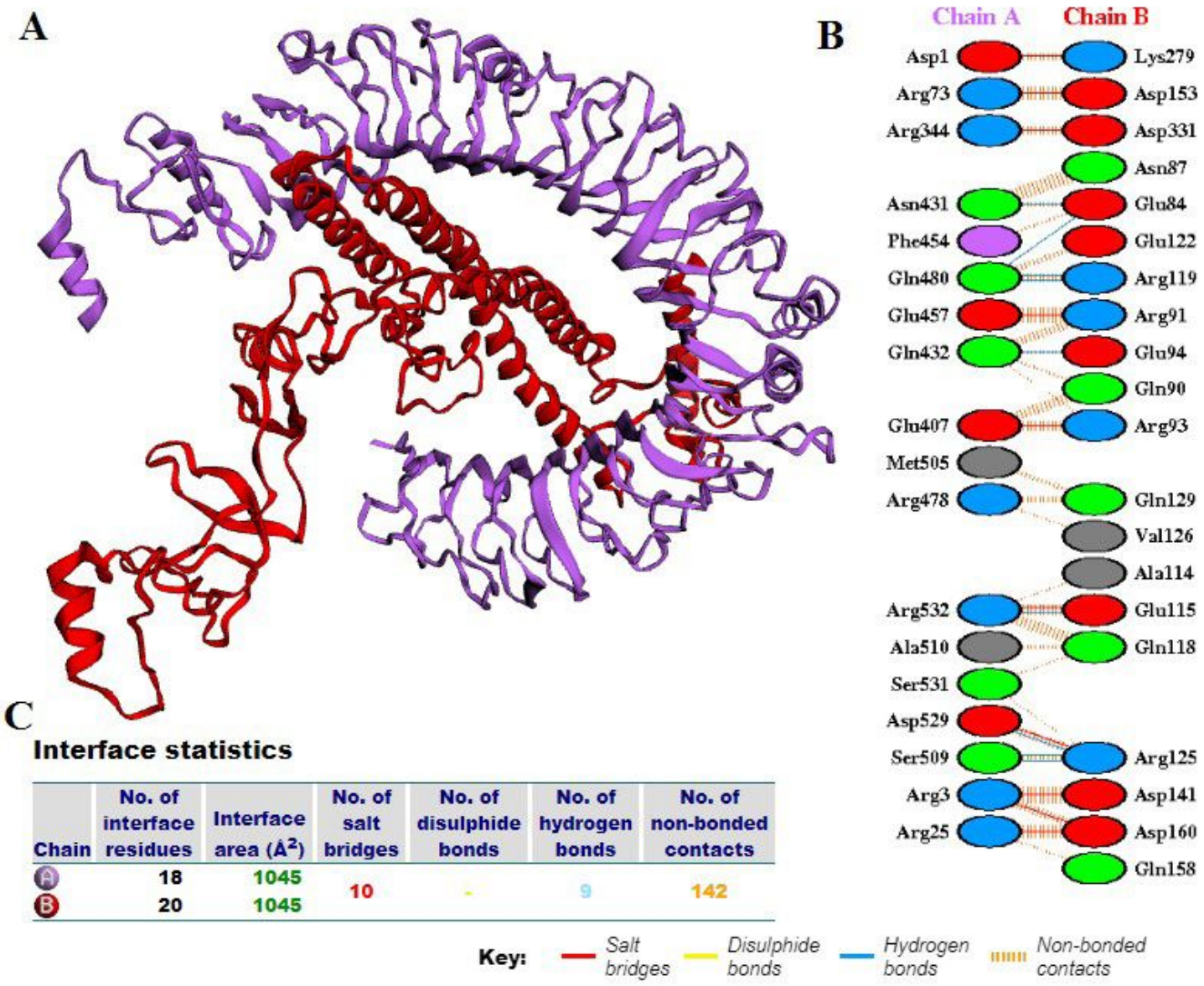

Figure 3

Vaccine construct-TLR5 docked complex: A) Figure obtained after molecular docking between Vaccine and TLR-3, red color showed the vaccine construct while purple color is representing TLR5. B) The interacting residues illustration between docked vaccine (chain B) and TLR5 (chain A) complex. C) Interface statics of vaccine construct-TLR5 docked complex.

Hydrogen bonds

	<----- ATOM 1 ----->					<----- ATOM 2 ----->				
	Atom no.	Atom name	Res name	Res no.	Chain		Atom no.	Atom name	Res name	Res no.	Chain	Distance
1.	4222	ND2	ASN	431	A	<-->	770	OE1	GLU	84	B	2.91
2.	4234	NE2	GLN	432	A	<-->	885	OE1	GLU	94	B	2.91
3.	4699	OE1	GLN	480	A	<-->	1116	NH1	ARG	119	B	3.08
4.	4699	OE1	GLN	480	A	<-->	1119	NH2	ARG	119	B	2.64
5.	4700	NE2	GLN	480	A	<-->	771	OE2	GLU	84	B	3.26
6.	4979	OG	SER	509	A	<-->	1181	NH1	ARG	125	B	3.25
7.	4979	OG	SER	509	A	<-->	1184	NH2	ARG	125	B	2.76
8.	5151	OD1	ASP	529	A	<-->	1181	NH1	ARG	125	B	2.79
9.	5178	NE	ARG	532	A	<-->	1073	OE1	GLU	115	B	2.90

Figure 4

Coordinates of hydrogen bonds between vaccine construct and receptor TLR5.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryTables110.pdf