**Designing Novel Subunit Vaccines against Herpes Simplex Virus-1 using Reverse Vaccinology Approach**

Bishajit Sarkar [1*] , Md. Asad Ullah [1]

[1]Department of Biotechnology and Genetic Engineering, Jahangirnagar University, Savar, Dhaka, Bangladesh.

Corresponding author: Bishajit Sarkar[*]

Email of the corresponding author: sarkarbishajit@gmail.com

ORCID iD of the corresponding author: https://orcid.org/0000-0001-8439-6994

**Abstract**

Herpes Simplex Virus (HSV) is an infectious virus that is responsible for various types of orofacial and genital infections. Two types of HSV viruses, HSV-1 and HSV-2, are the most dangerous HSV viruses. Every year, millions of people get infected with this menacing virus, however, no satisfactory treatment or vaccine has not yet been discovered to fight against HSV. Although there are some anti-viral therapies, however, studies have showed that such anti-viral therapies may also fail to provide good impact. In this study, a possible subunit vaccine against HSV-1, strain-17, was designed using the tools and reverse vaccinology and bioinformatics. Three potential antigenic envelope glycoproteins were selected from nine envelope glycoproteins, for possible vaccine construction. Potential epitopes capable of inducing high immunogenic response and at the same time have non-allergenicity and conservancy across other strains and species, were selected by some robust analysis, for vaccine construction. Finally, three possible vaccines were designed. Each of the vaccine construct differ from each other only in their adjuvant sequences and based on molecular docking analysis, one best vaccine construct was selected for molecular dynamics simulation study and in silico codon adaptation. The experiment showed that the selected best vaccine should be good candidate against HPV-1, strain-17. However, wet lab study should be conducted on the suggested vaccine(s) for confirming their potentiality, safety and efficacy.

Keywords: Herpes Simplex Virus-1, vaccine, reverse vaccinology, immunity, bioinformatics

## 1. Introduction

Herpes Simplex Virus (HSV) is a member of the herpesviridae family. HSV is a double stranded DNA containing virus with an icosapentahedral capsid. The HSV-1 and HSV-2 are the two most dangerous types of herpes. HSV-1 mainly causes the orofacial infections and HSV-2 usually causes genital infections [1, 2]. The HSV-1 virus is transmitted from individual to individual primarily by oral-oral contact and cause a disease which is known as cold sores. HSV-1 infection is wide-spread throughout the world and in many cases, the infection may remain latent or undetected throughout the whole life of the infected person. Sometimes, the infection may lead to much dangerous diseases like the encephalitis. On the other hand, the HSV-2 is mainly a sexually transmitted virus and therefore, it mainly causes the genital herpes. According to a study, the total number of new HSV-1 cases in 2012 was estimated to be 118 million [3, 4]. Moreover, the global incidence of HSV-2 was estimated to be about 417 million in 2012 [5]. Sometimes, the incidence of HSV infection leads to other diseases like the Alzheimer's disease and liver failure [6, 7]. Currently, the treatment of HSV is carried out using many antiviral medications like acyclovir, valacyclovir, famciclovir etc. However, in many cases, such antiviral medications have failed to provide complete resistant to HSV and reduce the mortality rate of the HSV infected patients. Moreover, studies have proved that the viral resistance of Herpes virus to such medications (like the acyclovir) may also occur and such incidence can make the treatments more complicated [8, 9, 10, 11]. Although researchers are working on to develop an effective vaccine against the Herpes virus like subunit vaccines, inactivated virion vaccines, genetically attenuated vaccines etc., to date, no satisfactory vaccine has been successfully entered the market [12]. In this study, the reverse vaccinology and bioinformatics approach was used to develop an effective Herpes vaccine against the HSV-1, strain-17 (**Figure 01**). Reverse vaccinology is defined as the process of vaccine

development where the novel antigens of a virus or organism are identified by analyzing the genomic information of that particular organism or a virus. In reverse vaccinology, various tools of bioinformatics are used for identifying the novel antigens useful for vaccine development by dissecting the genome as well as for studying the genetic makeup of a pathogen. This approach of vaccine development also helps the scientists to understand which antigenic segments of a virus or pathogen should be given emphasis during the vaccine development. This method is a quick easy and cost-effective way to design vaccine. Reverse vaccinology is successfully used for developing vaccines for many viruses like the Zika virus, Chikungunya virus etc. [13, 14].

## 2. Materials and methods:

The experiment focuses on developing of vaccines against the Human Herpes Simplex Virus- 1 (strain-17).

### 2.1. Strain Identification and Selection

The Herpes Simplex Virus- 1 (strain-17) was identified and selected by analyzing and reviewing different entries of the online server of National Center for Biotechnology Information or NCBI (https://www.ncbi.nlm.nih.gov/).

### 2.2. Retrieval of the Protein Sequences

Nine viral proteins: envelope glycoprotein M (accession number: P04288), envelope glycoprotein H (accession number: P06477), envelope glycoprotein E (accession number: P04488), envelope glycoprotein I (accession number: P06487), envelope glycoprotein B (accession number: P10211), envelope glycoprotein L (accession number: P10185), envelope glycoprotein D (accession number: Q69091), envelope glycoprotein K (accession number: P68331) and envelope glycoprotein C (accession number: P10228) were retrieved from the UniProt Knowledgebase (UniProtKB) tool of the online server UniProt (https://www.uniprot.org/).

### 2.3. Antigenicity Prediction and Physicochemical Property Analysis of the Protein Sequences

The antigenicity of the protein sequences were predicted by the online server, VaxiJen v2.0 (http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.htm), keeping the threshold at 0.4 and tumor model was used [15, 16, 17]. Only the antigenic proteins were selected for further analysis. The various physicochemical properties of the selected antigenic protein sequences were determined by ExPASy's online tool ProtParam (https://web.expasy.org/protparam/) [18].

**2.4. T-cell and B-cell Epitope Prediction**

The T-cell and B-cell epitopes of the selected protein sequences were predicted using online epitope prediction server Immune Epitope Database or IEDB (https://www.iedb.org/). The IEDB database contains huge amount of experimental data on T-cell epitopes as well as antibodies. These data are collected from various studies that are conducted on human, non-human primates and other animals. It is a server that allows robust analysis on various epitopes by exploiting various tools: population coverage, conservation across antigens and clusters with similar sequences [19]. The MHC class-I restricted CD8+ cytotoxic T-lymphocyte (CTL) epitopes of the selected sequences were obtained using NetMHCpan EL 4.0 prediction method for HLA-A*11-01 allele. The MHC class-II restricted CD4+ helper T-lymphocyte (HTL) epitopes were obtained for HLA DRB1*04-01 allele using Sturniolo prediction method. The top ten MHC class-I were selected based on their percentile scores and antigenicity scores (AS). Moreover, MHC class-II epitope prediction generated several epitopes with similar AS and percentile scores. However, 2 epitopes were taken from each of the top five categories. On the other hand, five random B-cell lymphocyte epitopes (BCL) were selected based on their length (the epitope sequences that had ten amino acids or above were selected) and obtained using Bipipered linear epitope prediction method, keeping the parameters default.

**2.5. Transmembrane Topology and Antigenicity Prediction of the Selected Epitopes**

The transmembrane topology of the selected epitopes were determined using the transmembrane topology of protein helices determinant, TMHMM v2.0 server (http://www.cbs.dtu.dk/services/TMHMM/). The server predicts whether the epitope would be transmembrane or it would remain inside or outside of the membrane [20]. The antigenicity of the

selected epitopes were predicted using VaxiJen v2.0 (http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.htm), using the tumor model and threshold of 0.4.

## 2.6. Allergenicity and Toxicity Prediction of the Epitopes

The allergenicity of the selected epitopes were predicted using two online tools, AllerTOP v2.0 (https://www.ddg-pharmfac.net/AllerTOP/) and AllergenFP v1.0 (http://ddg-pharmfac.net/AllergenFP/). However, the results predicted by AllerTOP were given priority since the server has better accuracy of 88.7% than AllergenFP server (87.9%) [21, 22]. The toxicity prediction of the selected epitopes were carried out using ToxinPred server (http://crdd.osdd.net/raghava/toxinpred/), using SVM (Swiss-Prot) based method, keeping all the parameters default.

## 2.7. Conservancy Prediction of the Selected Epitopes

The conservancy analysis of the selected epitopes were performed using the 'epitope conservancy analysis' tool of IEDB server (https://www.iedb.org/conservancy/) [19]. The sequence identity threshold was kept '>=50'. For the conservancy analysis of the selected epitopes of the Herpes Simplex Virus- 1 strain-17, the envelope glycoprotein E, envelope glycoprotein B and envelope glycoprotein D of Human Herpes Simplex Virus-1 strain-F and Human Herpes Simplex Virus-2 strain-HG52 were used for comparison along with the proteins of the Human Herpes Simplex Virus-1 strain- 17 itself (UniProt accession numbers: P89475, Q703F0, P04488, P08666, P06436, P10211, Q69467, Q05059 and Q69091). This ensures that, the epitope sequences are conserved across the species as well as strains. Based on the antigenicity, allergenicity, toxicity and conservancy analysis, the best ligands were selected for the further analysis and vaccine construction. The epitopes that showed antigenicity, non-allergenicity, non-toxicity and high

(more than 90%) conservancy and more than 50% minimum identity, were considered as the best epitopes. For B-cell epitope selection, only the antigenic and non-allergenic epitopes were taken for further analysis.

## 2.8. Cluster Analysis of the MHC Alleles

Cluster analysis of the MHC alleles helps to identify the alleles of the MHC class-I and class-II molecules that have similar binding specificities. The cluster analysis of the MHC alleles were carried out by online tool MHCcluster 2.0 (http://www.cbs.dtu.dk/services/MHCcluster/) [23]. During the analysis, the number of peptides to be included was kept 50,000, the number of bootstrap calculations were kept 100 and all the HLA supertype representatives (MHC class-I) and HLA-DR representatives (MHC class-II) were selected. For analyzing the MHC class-I alleles, the NetMHCpan-2.8 prediction method was used. The output of the server generated results in the form of MHC specificity tree and MHC specificity heat-map.

## 2.9. Generation of the 3D Structures of the Selected Epitopes

The 3D structures of the selected best epitopes were generated using online 3D generating tool PEP-FOLD3 (http://bioserv.rpbs.univ-paris-diderot.fr/services/PEP-FOLD3/). PEP-FOLD3 is a online tool for generating de novo peptide 3D structure [24, 25, 26].

## 2.10. Molecular Docking of the Selected Epitopes

The molecular docking of the selected epitopes with the MHC class-I and class-II proteins were carried out by online docking tool PatchDock (https://bioinfo3d.cs.tau.ac.il/PatchDock/php.php). PatchDock tool divides the Connolly dot surface representation of the molecules into concave, convex and flat patches using its various algorithms. After that the complementary patches are matched for generating potential candidate transformations. Next, each of the candidate

transformations is evaluated by a scoring function and later, an RMSD (root mean square deviation) clustering is applied to the candidate solutions for discarding the redundant solutions. The top score solutions are made the top ranked solutions by the server. After docking by PatchDock, the docking results were refined and re-scored by FireDock server (http://bioinfo3d.cs.tau.ac.il/FireDock/php.php). The FireDock server generates global energies upon the refinement of the best solutions from the PatchDock server and ranks them based on the generated global energies and the lowest global energy is always appreciable and preferred [27, 28, 29, 30]. The molecular docking experiments were carried out using the HLA-A*11-01 allele (PDB ID: 5WJL) and HLA DRB1*04-01 (PDB ID: 5JLZ) as receptors and the ligands were the best selected MHC-I and MHC-II epitopes, respectively. The receptors were downloaded from the RCSB- Protein Data Bank (https://www.rcsb.org/) server. The best results were visualized using Discovery Studio Visualizer [31].

## 2.11. Vaccine Construction

Three possible vaccines were constructed against the Herpes Simplex Virus-1, strain-17. The predicted CTL, HTL and BCL epitopes were conjugated together for constructing the vaccines. All the vaccines were generated maintaining the sequence: adjuvant, PADRE sequence, CTL epitopes, HTL epitopes and BCL epitopes. Three different adjuvant sequences were used for constructing three different vaccines: beta defensin, L7/L12 ribosomal protein and HABA protein (*M. tuberculosis*, accession number: AGV15514.1). Beta-defensin adjuvants induce the activation of the toll like receptors (TLRs): 1, 2 and 4, where beta-defensin acts as agonist. The L7/L12 ribosomal protein and HABA protein activate the TLR-4. During the vaccine construction, various

linkers were used: EAAAK linkers were used to conjugate the adjuvant and PADRE sequence, GGGS linkers were used to attach the PADRE sequence with the CTL epitopes and the CTL epitopes with the other CTL epitopes, GPGPG linkers were used to connect the CTL with HTL epitopes and also the HTL epitopes among themselves. The KK linkers were used for conjugating the HTL and BCL epitopes as well as the BLC epitopes among themselves [32, 33, 34, 35, 36, 37, 38]. Studies have proved that the PADRE sequence improves the CTL response of the vaccines that contain it [39]. Total three vaccines were constructed in the experiment.

## 2.12. Antigenicity, Allergenicity and Physicochemical Property Analysis

The antigenicity of the constructed vaccines were determined by the online server VaxiJen v2.0 (http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.htm). The threshold of the prediction was kept at 0.4. AlgPred (http://crdd.osdd.net/raghava/algpred/) and AllerTop v2.0 (https://www.ddg-pharmfac.net/AllerTOP/) were used for the prediction of the allergenicity of the vaccine constructs. The AlgPred server predicts the possible allergens based on similarity of known epitope of any of the known region of the protein [40]. MEME/MAST motif prediction approach is used in the allergenicity prediction of the vaccines by AlgPred. Moreover, various physicochemical properties of the vaccines were examined by the online server ProtParam (https://web.expasy.org/protparam/).

## 2.13. Secondary and Tertiary Structure Prediction of the Vaccine Constructs

The secondary structures of the vaccine constructs were generated using online tool PRISPRED (http://bioinf.cs.ucl.ac.uk/psipred/). PRISPRED is a simple secondary structure generator which can be used to predict the transmembrane topology, transmembrane helix, fold and domain

recognition etc. along with the secondary structure prediction [41, 42]. The PRISPRED 4.0 prediction method was used to predict the secondary structures of the vaccine constructs. The β-sheet structure of the vaccines were determined by another online tool, NetTurnP v1.0 (http://www.cbs.dtu.dk/services/NetTurnP/) [43]. The tertiary or 3D structures of the vaccines were generated using online tool RaptorX (http://raptorx.uchicago.edu/) server. The server is a fully annotated tool for the prediction of protein structure, the property and contact prediction, sequence alignment etc. [44, 45, 46].

## 2.14. 3D Structure Refinement and Validation

The 3D structures of the constructed vaccines were refined using online refinement tool, 3Drefine (http://sysbio.rnet.missouri.edu/3Drefine/). The server is a quick, easy and efficient tool for protein structure refinement [47]. For each of the vaccine, the refined model 1 was downloaded for validation. The refined vaccine proteins were then validated by analyzing the Ramachandran plots which were generated using the online tool, PROCHECK (https://servicesn.mbi.ucla.edu/PROCHECK/) [48, 49].

## 2.15. Vaccine Protein Disulfide Engineering

The vaccine protein disulfide engineering was carried out by online tool Disulfide by Design 2 v12.2 (http://cptweb.cpt.wayne.edu/DbD2/). The server predicts the possible sites within a protein structure which have the greater possibility of undergoing disulfide bond formation [50]. When engineering the disulfide bonds, the intra-chain, inter-chain and $C_\beta$ for glycine residue were selected. The $\chi_3$ Angle was kept -87° or +97° ± 5 and $C_\alpha$-$C_\beta$-$S_\gamma$ Angle was kept 114.6° ±10.

## 2.16. Protein-Protein Docking

In protein-protein docking, the constructed Herpes Simplex Virus-1, strain-17 virus vaccines were analyzed by docking against various MHC alleles and toll like receptor (TLR). One best vaccine was selected based on their performances in the docking experiment. When viral infections occur, the viral particles are recognized by the MHC complex as antigens. The various segments of the MHC molecules are encoded by different alleles. For this reason, the vaccines should have good binding affinity with these MHC portions that are encoded by different alleles [51]. All the vaccine constructs were docked against the selected MHC alleles to test their binding affinity. In this experiment, the vaccines constructs were docked against DRB1*0101 (PDB ID: 2FSE), DRB3*0202 (PDB ID: 1A6A), DRB5*0101 (PDB ID: 1H15), DRB3*0101 (PDB ID: 2Q6W), DRB1*0401 (PDB ID: 2SEB), and DRB1*0301 (PDB ID: 3C5J) alleles. Moreover, studies have proved that TLR-8 that are present on the immune cells, are responsible for mediating the immune responses against the RNA viruses and TLR-3 of the immune cells mediates immune responses against the DNA viruses [52, 53]. The Herpes virus is a DNA virus 9540. For this reason, the vaccine constructs of Herpes virus were also docked against TLR-3 (PDB ID: 2A0Z). The protein-protein docking was carried out using various online docking tools. The docking was carried out three times by three different online servers to improve the accuracy of the docking. First, the docking was carried out by ClusPro 2.0 (https://cluspro.bu.edu/login.php). The server ranks the clusters of docked complexes based on their center and lowest energy scores. However, these scores do not represent the actual binding affinity of the proteins with their targets [55, 56, 57]. The bonding affinity ($\Delta G$ in kcal mol$^{-1}$) of docked complexes were generated by PRODIGY tool of HADDOCK webserver (https://haddock.science.uu.nl/). The lower the binding energy generated by the server, the higher the binding binding affinity [58, 59, 60]. Moreover, the docking was again performed by PatchDock (https://bioinfo3d.cs.tau.ac.il/PatchDock/php.php) and later

refined and re-scored by FireDock server (http://bioinfo3d.cs.tau.ac.il/FireDock/php.php). The FireDock server ranks the docked complexes based on their global energy and the lower the score, the better the result. Later, the docking was performed using HawkDock server (http://cadd.zju.edu.cn/hawkdock/). The Molecular Mechanics/Generalized Born Surface Area (MM-GBSA) study was also carried out using HawkDock server. According to the server, the lower score and lower energy corresponds to better scores [61, 62, 63, 64]. The HawkDock server generates several models of docked complex and ranks them by assigning HawkDock scores in the ascending order. For each of the vaccines and their respective targets, the score of model 1 was taken for analysis. Furthermore, the model 1 of every complex was analyzed for MM-GBSA study. From the docking experiment, one best vaccine was selected for further analysis. The docked structures were visualized by PyMol tool [65].

### 2.17. Molecular Dynamic Simulation

The molecular dynamics simulation study was performed on the best selected vaccine, HV-1. The study was carried out by the online server iMODS (http://imods.chaconlab.org/). iMODS can be used efficiently to investigate the structural dynamics of the protein complexes, since it is a quick, easy and user-friendly server. The server provides the values of deformability, B-factor (mobility profiles), eigenvalues, variance, co-variance map and elastic network, for a protein or docked protein-protein complex. The deformability depends on the ability to deform at each of its amino acid. The eigenvalue is related with the energy that is required to deform the given structure and the lower eigenvalue corresponds to the easier the deformability of the complex. The eigenvalue also represents the motion stiffness of the protein complex. The server is a fast and easy tool for

determining and measuring the protein flexibility [66, 67, 68, 69, 70]. For analysing the molecular dynamics simulation of the HV-1-TLR-3 docked complex was used. The docked PDB files were uploaded to the iMODS server and the results were displayed keeping all the parameters as default.

### 2.18. Codon Adaptation and In Silico Cloning

The codon adaptation and in silico cloning were carried out for the best selected vaccine protein, HV-1. For conducting these experiments, the vaccine protein was reverse transcribed to the possible DNA sequences. The DNA sequence should encode the target vaccine protein. Later, the reverse transcribed DNA sequences were adapted according to the desired organism, so that the cellular mechanisms of that specific organism could use the codons of the adapted DNA sequences efficiently and provide better production of the desired product. Codon adaptation is a necessary step of in silico cloning since the same amino acid can be encoded by different codons in different organisms, a phenomenon which is known as codon biasness. Moreover, the cellular mechanisms of an organism may be different from another organism and a codon for a specific amino acid may not work in another organism. For this reason, codon adaptation step is performed that can predict the suitable codon that can encode a specific amino acid in a specific organism [71, 72]. The codon adaptation of the selected vaccine protein was carried out by the Java Codon Adaptation Tool or JCat server (http://www.jcat.de/) [71]. Eukaryotic *E. coli* strain K12 was selected and rho-independent transcription terminators, prokaryotic ribosome binding sites and SgrA1 and SphI cleavage sites of restriction enzymes, were avoided. In the JCat server, the protein sequences were reverse translated to the optimized possible DNA sequences. The optimized DNA sequences were taken and SgrA1 and SphI restriction sites were conjugated at the N-terminal and C-terminal sites,

respectively. Next, the SnapGene [73] restriction cloning module was used to insert the new

adapted DNA sequences between SgrA1 and SphI of pET-19b vector.

## 3. Results

## 3.1. Identification, Selection and Retrieval of Viral Protein Sequences

The Herpes Simplex Virus, strain-17 was identified and selected from the NCBI (https://www.ncbi.nlm.nih.gov/). Nine envelope proteins from the viral structure was selected for the possible vaccine construction. These proteins were: envelope glycoprotein M (accession number: P04288), envelope glycoprotein H (accession number: P06477), envelope glycoprotein E (accession number: P04488), envelope glycoprotein I (accession number: P06487), envelope glycoprotein B (accession number: P10211), envelope glycoprotein L (accession number: P10185), envelope glycoprotein D (accession number: Q69091), envelope glycoprotein K (accession number: P68331) and envelope glycoprotein C (accession number: P10228). The protein sequences were retrieved from the online server, UniProt (https://www.uniprot.org/). The protein sequences in fasta format:

>sp|P04288|GM_HHV11 Envelope glycoprotein M OS=Human herpesvirus 1 (strain 17) OX=10299 GN=gM PE=1 SV=1

MGRPAPRGSPDSAPPTKGMTGARTAWWVWCVQVATFVVSAVCVTGLLVLASVFRARF
PCFYATASSYAGVNSTAEVRGGVAVPLRLDTQSLVGTYVITAVLLLAVAVYAVVGAVT
SRYDRALDAGRRLAAARMAMPHATLIAGNVCSWLLQITVLLLAHRISQLAHLVYVLHF
ACLVYFAAHFCTRGVLSGTYLRQVHGLMELAPTHHRVVGPARAVLTNALLLGVFLCTA
DAAVSLNTIAAFNFNFSAPGMLICLTVLFAILVVSLLLVVEGVLCHYVRVLVGPHLGAV
AATGIVGLACEHYYTNGYYVVETQWPGAQTGVRVALALVAAFALGMAVLRCTRAYL
YHRRHHTKFFMRMRDTRHRAHSALKRVRSSMRGSRDGRHRPAPGSPPGIPEYAEDPYAI
SYGGQLDRYGDSDGEPIYDEVADDQTDVLYAKIQHPRHLPDDDPIYDTVGGYDPEPAED
PVYSTVRRW

>sp|P06477|GH_HHV11 Envelope glycoprotein H OS=Human herpesvirus 1 (strain 17) OX=10299 GN=gH PE=1 SV=1

MGNGLWFVGVIILGVAWGQVHDWTEQTDPWFLDGLGMDRMYWRDTNTGRLWLPNT

PDPQKPPRGFLAPPDELNLTTASLPLLRWYEERFCFVLVTTAEFPRDPGQLLYIPKTYLLG

RPPNASLPAPTTVEPTAQPPPSVAPLKGLLHNPAASVLLRSRAWVTFSAVPDPEALTFPR

GDNVATASHPSGPRDTPPPRPPVGARRHPTTELDITHLHNASTTWLATRGLLRSPGRYV

YFSPSASTWPVGIWTTGELVLGCDAALVRARYGREFMGLVISMHDSPPVEVMVVPAGQ

TLDRVGDPADENPPGALPGPPGGPRYRVFVLGSLTRADNGSALDALRRVGGYPEEGTN

YAQFLSRAYAEFFSGDAGAEQGPRPPLFWRLTGLLATSGFAFVNAAHANGAVCLSDLL

GFLAHSRALAGLAARGAAGCAADSVFFNVSVLDPTARLQLEARLQHLVAEILEREQSLA

LHALGYQLAFVLDSPSAYDAVAPSAAHLIDALYAEFLGGRVLTTPVVHRALFYASAVLR

QPFLAGVPSAVQRERARRSLLIASALCTSDVAAATNADLRTALARADHQKTLFWLPDHF

SPCAASLRFDLDESVFILDALAQATRSETPVEVLAQQTHGLASTLTRWAHYNALIRAFVP

EASHRCGGQSANVEPRILVPITHNASYVVTHSPLPRGIGYKLTGVDVRRPLFLTYLTATC

EGSTRDIESKRLVRTQNQRDLGLVGAVFMRYTPAGEVMSVLLVDTDNTQQQIAAGPTE

GAPSVFSSDVPSTALLLFPNGTVIHLLAFDTQPVAAIAPGFLAASALGVVMITAALAGILK

VLRTSVPFFWRRE


>sp|P04488|GE_HHV11 Envelope glycoprotein E OS=Human herpesvirus 1 (strain 17) OX=10299 GN=gE PE=1 SV=1

MDRGAVVGFLLGVCVVSCLAGTPKTSWRRVSVGEDVSLLPAPGPTGRGPTQKLLWAV

EPLDGCGPLHPSWVSLMPPKQVPETVVDAACMRAPVPLAMAYAPPAPSATGGLRTDFV

WQERAAVVNRSLVIHGVRETDSGLYTLSVGDIKDPARQVASVVLVVQPAPVPTPPPTPA

DYDEDDNDEGEDESLAGTPASGTPRLPPPPAPPRSWPSAPEVSHVRGVTVRMETPEAILF

SPGETFSTNVSIHAIAHDDQTYSMDVVWLRFDVPTSCAEMRIYESCLYHPQLPECLSPAD

APCAASTWTSRLAVRSYAGCSRTNPPPRCSAEAHMEPVPGLAWQAASVNLEFRDASPQ

HSGLYLCVVYVNDHIHAWGHITISTAAQYRNAVVEQPLPQRGADLAEPTHPHVGAPPH

APPTHGALRLGAVMGAALLLSALGLSVWACMTCWRRRAWRAVKSRASGKGPTYIRV

ADSELYADWSSDSEGERDQVPWLAPPERPDSPSTNGSGFEILSPTAPSVYPRSDGHQSRR

QLTTFGSGRPDRRYSQASDSSVFW


>sp|P06487|GI_HHV11 Envelope glycoprotein I OS=Human herpesvirus 1 (strain 17) OX=10299 GN=gI PE=1 SV=1

MPCRPLQGLVLVGLWVCATSLVVRGPTVSLVSNSFVDAGALGPDGVVEEDLLILGELRF

VGDQVPHTTYYDGGVELWHYPMGHKCPRVVHVVTVTACPRRPAVAFALCRATDSTHS

PAYPTLELNLAQQPLLRVQRATRDYAGVYVLRVWVGDAPNASLFVLGMAIAAEGTLA

YNGSAYGSCDPKLLPSSAPRLAPASVYQPAPNQASTPSTTTSTPSTTIPAPSTTIPAPQAST

TPFPTGDPKPQPPGVNHEPPSNATRATRDSRYALTVTQIIQIAIPASIIALVFLGSCICFIHRC

QRRYRRSRRPIYSPQMPTGISCAVNEAAMARLGAELKSHPSTPPKSRRRSSRTPMPSLTAI

AEESEPAGAAGLPTPPVDPTTPTPTPPLLV


>sp|P10211|GB_HHV11 Envelope glycoprotein B OS=Human herpesvirus 1 (strain 17) OX=10299 GN=gB PE=1 SV=1

MRQGAPARGRRWFVVWALLGLTLGVLVASAAPSSPGTPGVAAATQAANGGPATPAPP

APGAPPTGDPKPKKNRKPKPPKPPRPAGDNATVAAGHATLREHLRDIKAENTDANFYV

CPPPTGATVVQFEQPRRCPTRPEGQNYTEGIAVVFKENIAPYKFKATMYYKDVTVSQV

WFGHRYSQFMGIFEDRAPVPFEEVIDKINAKGVCRSTAKYVRNNLETTAFHRDDHETD

MELKPANAATRTSRGWHTTDLKYNPSRVEAFHRYGTTVNCIVEEVDARSVYPYDEFVL

ATGDFVYMSPFYGYREGSHTEHTSYAADRFKQVDGFYARDLTTKARATAPTTRNLLTT

PKFTVAWDWVPKRPSVCTMTKWQEVDEMLRSEYGGSFRFSSDAISTTFTTNLTEYPLSR

VDLGDCIGKDARDAMDRIFARRYNATHIKVGQPQYYLANGGFLIAYQPLLSNTLAELYV

REHLREQSRKPPNPTPPPPGASANASVERIKTTSSIEFARLQFTYNHIQRHVNDMLGRVAI

AWCELQNHELTLWNEARKLNPNAIASATVGRRVSARMLGDVMAVSTCVPVAADNVIV

QNSMRISSRPGACYSRPLVSFRYEDQGPLVEGQLGENNELRLTRDAIEPCTVGHRRYFTF

GGGYVYFEEYAYSHQLSRADITTVSTFIDLNITMLEDHEFVPLEVYTRHEIKDSGLLDYT

EVQRRNQLHDLRFADIDTVIHADANAAMFAGLGAFFEGMGDLGRAVGKVVMGIVGGV

VSAVSGVSSFMSNPFGALAVGLLVLAGLAAAFFAFRYVMRLQSNPMKALYPLTTKELK

NPTNPDASGEGEEGGDFDEAKLAEAREMIRYMALVSAMERTEHKAKKKGTSALLSAKV

TDMVMRKRRNTNYTQVPNKDGDADEDDL


>sp|P10185|GL_HHV11 Envelope glycoprotein L OS=Human herpesvirus 1 (strain 17) OX=10299 GN=gL PE=1 SV=1

MGILGWVGLIAVGVLCVRGGLPSTEYVIRSRVAREVGDILKVPCVPLPSDDLDWRYETP

SAINYALIDGIFLRYHCPGLDTVLWDRHAQKAYWVNPFLFVAGFLEDLSYPAFPANTQE

TETRLALYKEIRQALDSRKQAASHTPVKAGCVNFDYSRTRRCVGRQDLGPTNGTSGRTP

VLPPDDEAGLQPKPLTTPPPIIATSDPTPRRDAATKSRRRRPHSRRL

>sp|Q69091|GD_HHV11 Envelope glycoprotein D OS=Human herpesvirus 1 (strain 17) OX=10299 GN=gD PE=1 SV=1

MGGAAARLGAVILFVVIVGLHGVRSKYALVDASLKMADPNRFRGKDLPVLDQLTDPPG

VRRVYHIQAGLPDPFQPPSLPITVYYAVLERACRSVLLNAPSEAPQIVRGASEDVRKQPY

NLTIAWFRMGGNCAIPITVMEYTECSYNKSLGACPIRTQPRWNYYDSFSAVSEDNLGFL

MHAPAFETAGTYLRLVKINDWTEITQFILEHRAKGSCKYALPLRIPPSACLSPQAYQQGV

TVDSIGMLPRFIPENQRTVAVYSLKIAGWHGPKAPYTSTLLPPELSETPNATQPELAPEDP

EDSALLEDPVGTVAPQIPPNWHIPSIQDAATPYHPPATPNNMGLIAGAVGGSLLAALVIC

GIVYWMRRHTQKAPKRIRLPHIREDDQPSSHQPLFY


>sp|P68331|GK_HHV11 Envelope glycoprotein K OS=Human herpesvirus 1 (strain 17) OX=10299 GN=gK PE=1 SV=1

MLAVRSLQHLSTVVLITAYGLVLVWYTVFGASPLHRCIYAVRPTGTNNDTALVWMKM

NQTLLFLGAPTHPPNGGWRNHAHICYANLIAGRVVPFQVPPDAMNRRIMNVHEAVNCL

ETLWYTRVRLVVVGWFLYLAFVALHQRRCMFGVVSPAHKMVAPATYLLNYAGRIVSS

VFLQYPYTKITRLLCELSVQRQNLVQLFETDPVTFLYHRPAIGVIVGCELMLRFVAVGLI

VGTAFISRGACAITYPLFLTITTWCFVSTIGLTELYCILRRGPAPKNADKAAAPGRSKGLS

GVCGRCCSIILSGIAVRLCYIAVVAGVVLVALHYEQEIQRRLFDV


>sp|P10228|GC_HHV11 Envelope glycoprotein C OS=Human herpesvirus 1 (strain 17) OX=10299 GN=gC PE=1 SV=1

MAPGRVGLAVVLWSLLWLGAGVSGGSETASTGPTITAGAVTNASEAPTSGSPGSAASPE

VTPTSTPNPNNVTQNKTTPTEPASPPTTPKPTSTPKSPPTSTPDPKPKNNTTPAKSGRPTKP

PGPVWCDRRDPLARYGSRVQIRCRFRNSTRMEFRLQIWRYSMGPSPPIAPAPDLEEVLTN

ITAPPGGLLVYDSAPNLTDPHVLWAEGAGPGADPPLYSVTGPLPTQRLIIGEVTPATQGM

YYLAWGRMDSPHEYGTWVRVRMFRPPSLTLQPHAVMEGQPFKATCTAAAYYPRNPVE

FVWFEDDHQVFNPGQIDTQTHEHPDGFTTVSTVTSEAVGGQVPPRTFTCQMTWHRDSV

TFSRRNATGLALVLPRPTITMEFGVRIVVCTAGCVPEGVTFAWFLGDDPSPAAKSAVTA

QESCDHPGLATVRSTLPISYDYSEYICRLTGYPAGIPVLEHHGSHQPPPRDPTERQVIEAIE

WVGIGIGVLAAGVLVVTAIVYVVRTSQSRQRHRR

**3.2. Antigenicity Prediction and Physicochemical Property Analysis of the Protein Sequences**

From the selected nine proteins, only the antigenic proteins were selected for vaccine construction. VaxiJen v2.0 (http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.htm) server was used for antigenicity determination. Among the nine protein, envelope glycoprotein E, envelope glycoprotein B and envelope glycoprotein D were determined as possible antigenic proteins by the server (**Table 01**). The physicochemical property analysis was conducted on these three proteins.

The physicochemical study revealed that envelope glycoprotein B had the highest extinction co-efficient of 105255 $M^{-1}$ $cm^{-1}$ and lowest GRAVY value of -0.403. However, all the three proteins showed instability and all of them had half-lives of 30 hours in the mammalian cell culture system. The results of the physicochemical property analysis are listed in **Table 02.**

**Table 01**. The antigenicity determination of the nine selected proteins.

| Name of the protein | Antigenicity (threshold = 0.4; tumor model) |
|---|---|
| Envelope glycoprotein M | Non-antigenic |
| Envelope glycoprotein H | Non-antigenic |
| Envelope glycoprotein E | Antigenic |
| Envelope glycoprotein I | Non-antigenic |
| Envelope glycoprotein B | Antigenic |
| Envelope glycoprotein L | Non-antigenic |
| Envelope glycoprotein D | Antigenic |
| Envelope glycoprotein K | Non-antigenic |
| Envelope glycoprotein C | Non-antigenic |

**Table 02**. The antigenicity and physicochemical property analysis of the selected viral proteins.

| Name of the protein sequence | Total amino acids | Molecular weight | Theoretical pI | Ext. coefficient (in M$^{-1}$ cm$^{-1}$) | Est. half-life (in mammalian cell) | Instability index | Aliphatic index | Grand average of hydropathicity (GRAVY) |
|---|---|---|---|---|---|---|---|---|
| Envelope glycoprotein E | 550 | 59093.64 | 5.74 | 104110 | 30 hours | 58.52 (unstable) | 74.15 | -0.255 |
| Envelope glycoprotein B | 904 | 100292.44 | 8.30 | 105255 | 30 hours | 40.03 (unstable) | 70.83 | -0.403 |
| Envelope glycoprotein D | 394 | 43346.88 | 7.64 | 58705 | 30 hours | 61.11 (unstable) | 89.42 | -0.143 |

### 3.3. T-cell and B-cell Epitope Prediction and Topology Determination of the Epitopes

The T-cell epitopes of MHC class-I of the three proteins were determined by NetMHCpan EL 4.0 prediction method of the IEDB (https://www.iedb.org/) server keeping the sequence length 09. The server yielded over 100 such epitopes. However, based on analyzing the best antigenicity scores (AS) and percentile scores for each epitope, the top ten potential epitopes were selected randomly for antigenicity, allergenicity, toxicity and conservancy tests. The server ranks the predicted epitopes based on the ascending order of percentile scores. The lower the percentile, the better the binding affinity. The T-cell epitopes of MHC class-II (HLA DRB1*04-01 allele) of the proteins were also determined by IEDB (https://www.iedb.org/) server, where the Sturniolo prediction method was used. Among the hundreds of epitopes generated by the server, best ten epitopes were selected for further analysis. Moreover, the B-cell epitopes of the proteins were selected using Bipipered linear epitope prediction method of the IEDB (https://www.iedb.org/) server and epitopes were selected based on their higher lengths (**Figure 02**). The topology of the selected epitopes were determined by TMHMM v2.0 server (http://www.cbs.dtu.dk/services/TMHMM/). **Table 03** and **Table 04** list the potential T-cell epitopes of envelope protein E, **Table 05** and **Table 06** list the potential T-cell epitopes of envelope glycoprotein B, **Table 07** and **Table 08** list the potential T-cell epitopes of envelope protein D and **Table 09** list the potential B-cell epitopes with their respective topologies.

**Table 03**. MHC class-I epitope prediction and topology, antigenicity, allergenicity, toxicity and conservancy analysis of the epitopes of envelope glycoprotein E.

| Envelope glycoprotein E | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Epitope | Start | End | Topology | AS | Percentile scores | Antigenicity (tumor model, threshold= 0.4) | Allergenicity | Toxicity | Conservancy | Minimum identity |
| YSMDVVWLR | 256 | 264 | Inside | 0.676 | 0.16 | Non-antigen | allergen | Non-toxic | 100% | 77.78% |
| YTLSVGDIK | 140 | 148 | Inside | 0.566 | 0.26 | Antigen | Non-allergen | Non-toxic | 100% | 55.56% |
| ITISTAAQY | 373 | 381 | Inside | 0.539 | 0.29 | Antigen | Non-allergen | Non-toxic | 100% | 77.78% |
| TISTAAQYR | 374 | 382 | Inside | 0.412 | 0.48 | Antigen | Allergen | Non-toxic | 100% | 66.67% |
| VSCLAGTPK | 16 | 24 | Inside | 0.362 | 0.57 | Antigen | Allergen | Non-toxic | 100% | 66.67% |
| RTDFVWQER | 111 | 119 | Inside | 0.335 | 0.63 | Non-antigen | Allergen | Non-toxic | 100% | 33.33% |
| RIYESCLYH | 275 | 283 | Inside | 0.335 | 0.63 | Non-antigen | Allergen | Non-toxic | 100% | 88.89% |
| VVEQPLPQR | 385 | 393 | Outside | 0.316 | 0.69 | Non-antigen | Non-allergen | Non-toxic | 100% | 88.89% |
| AVKSRASGK | 451 | 459 | Inside | 0.308 | 0.72 | Antigen | Non-allergen | Non-toxic | 100% | 77.78% |
| GTPKTSWRR | 21 | 29 | Inside | 0.268 | 0.84 | Non-antigen | Non-allergen | Non-toxic | 100% | 55.56% |

**Table 04**. MHC class-II epitope prediction and topology, antigenicity, allergenicity, toxicity and conservancy analysis of the epitopes of envelope protein E.

| Epitope | Start | End | Topology | AS | Percentile scores | Antigenicity (tumor model, threshold= 0.4) | Allergenicity | Toxicity | Conservancy | Minimum identity |
|---|---|---|---|---|---|---|---|---|---|---|
| Envelope glycoprotein E | | | | | | | | | | |
| VTVRMETPEAIL | 222 | 233 | Inside | 4.00 | 0.93 | Antigen | Allergen | Non-toxic | 100% | 83.33% |
| VRMETPEAILFS | 224 | 235 | Outside | 4.00 | 0.93 | Antigen | Non-allergen | Non-toxic | 100% | 75.00% |
| WLRFDVPTSCAE | 262 | 273 | Outside | 3.30 | 2.40 | Antigen | Allergen | Non-toxic | 100% | 75.00% |
| LRFDVPTSCAEM | 263 | 274 | Inside | 3.30 | 2.40 | Antigen | Non-allergen | Non-toxic | 100% | 75.00% |
| SWVSLMPPKQVP | 69 | 80 | Outside | 2.90 | 3.60 | Non-antigen | Allergen | Non-toxic | 100% | 58.33% |
| PSWVSLMPPKQV | 68 | 79 | Outside | 2.90 | 3.60 | Antigen | Non-allergen | Non-toxic | 100% | 67.66% |
| SELYADWSSDSE | 469 | 480 | Outside | 2.80 | 3.90 | Non-antigen | allergen | Non-toxic | 100% | 100% |
| LYADWSSDSEGE | 471 | 482 | Outside | 2.80 | 3.90 | Non-antigen | allergen | Non-toxic | 100% | 100% |
| YRNAVVEQPLPQ | 381 | 392 | Outside | 2.50 | 5.30 | Non-antigen | allergen | Non-toxic | 100% | 100% |
| QYRNAVVEQPLP | 380 | 391 | Inside | 2.50 | 5.30 | Antigen | Allergen | Non-toxic | 100% | 91.67% |

**Table 05**. MHC class-I epitope prediction and topology, antigenicity, allergenicity, toxicity and conservancy analysis of the epitopes of envelope protein B.

| Envelope glycoprotein B | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Epitope | Start | End | Topology | AS | Percentile scores | Antigenicity (tumor model, threshold= 0.4) | Allergen icity | Toxicity | Conser vancy | Minimu m identity |
| KVTDMVM RK | 875 | 883 | Inside | 0.962 | 0.01 | Antigen | Non-allergen | Non-toxic | 100% | 77.78% |
| GTSALLSA K | 867 | 875 | Outside | 0.886 | 0.02 | Non-antigen | Non-allergen | Non-toxic | 100% | 88.89% |
| KALYPLTT K | 807 | 815 | Inside | 0.868 | 0.03 | Non-antigen | Non-allergen | Non-toxic | 100% | 100% |
| AIASATVGR | 549 | 557 | Inside | 0.753 | 0.10 | Non-antigen | Non-allergen | Non-toxic | 100% | 100% |
| TVAWDWV PK | 351 | 359 | Outside | 0.747 | 0.10 | Non-antigen | Non-allergen | Non-toxic | 100% | 100% |
| SAMERTEH K | 854 | 862 | Inside | 0.695 | 0.15 | Antigen | Allergen | Non-toxic | 100% | 100% |
| YAYSHQLS R | 653 | 661 | Inside | 0.682 | 0.15 | Antigen | Non-allergen | Non-toxic | 100% | 100% |
| FTFGGGYV Y | 641 | 649 | Outside | 0.661 | 0.17 | Non-antigen | Allergen | Non-toxic | 100% | 88.89% |
| ASANASVE R | 486 | 494 | Inside | 0.655 | 0.18 | Antigen | Non-allergen | Non-toxic | 100% | 88.89% |
| TTSSIEFAR | 497 | 505 | Inside | 0.554 | 0.27 | Non-antigen | Non-allergen | Non-toxic | 100% | 100% |

**Table 06**. MHC class-II epitope prediction and topology, antigenicity, allergenicity, toxicity and conservancy analysis of the epitopes of envelope protein B.

| Envelope glycoprotein B | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Epitope | Start | End | Topology | AS | Percentile scores | Antigenicity (tumor model, threshold= 0.4) | Allergenicity | Toxicity | Conservancy | Minimum identity |
| YYKDVTVSQVWF | 164 | 175 | Inside | 4.78 | 0.22 | Non-antigen | Non-allergen | Non-toxic | 100% | 100% |
| TMYYKDVTVSQV | 162 | 173 | Inside | 4.78 | 0.22 | Antigen | Non-allergen | Non-toxic | 100% | 100% |
| YARDLTTKARAT | 326 | 337 | Inside | 3.90 | 1.10 | Antigen | Non-allergen | Non-toxic | 100% | 100% |
| FYARDLTTKARA | 325 | 336 | Inside | 3.90 | 1.10 | Antigen | Allergen | Non-toxic | 100% | 100% |
| FARLQFTYNHIQ | 503 | 514 | Inside | 3.50 | 1.90 | Non-antigen | Non-allergen | Non-toxic | 100% | 100% |
| ARLQFTYNHIQR | 504 | 515 | Inside | 3.50 | 1.90 | Non-antigen | Allergen | Non-toxic | 100% | 100% |
| GGSFRFSSDAIS | 381 | 392 | Outside | 3.40 | 2.20 | Non-antigen | Allergen | Non-toxic | 100% | 100% |
| FRFSSDAISTTF | 384 | 395 | Inside | 3.40 | 2.20 | Non-antigen | Allergen | Non-toxic | 100% | 100% |

| VMRLQSNP MKAL | 798 | 809 | Inside | 3.30 | 2.40 | Antigen | Non-allergen | Non-toxic | 100% | 75% |
|---|---|---|---|---|---|---|---|---|---|---|
| RYVMRLQS NPMK | 796 | 807 | Inside | 3.30 | 2.40 | Antigen | Allergen | Non-toxic | 100% | 75% |

**Table 07**. MHC class-I epitope prediction and topology, antigenicity, allergenicity, toxicity and conservancy analysis of the epitopes of envelope protein D.

| Envelope glycoprotein D | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Epitope | Start | End | Topology | AS | Percentile scores | Antigenicity (tumor model, threshold= 0.4) | Allergenicity | Toxicity | Conservancy | Minimum identity |
| RTVAVYSLK | 254 | 262 | Inside | 0.885 | 0.02 | Non-antigen | Non-allergen | Non-toxic | 100% | 88.89% |
| TVYYAVLER | 81 | 89 | Inside | 0.842 | 0.05 | Non-antigen | Allergen | Non-toxic | 100% | 100% |
| RTQPRWNYY | 155 | 163 | Inside | 0.542 | 0.29 | Antigen | Allergen | Non-toxic | 100% | 88.89% |
| KMADPNRFR | 35 | 43 | Inside | 0.489 | 0.34 | Antigen | Non-allergen | Non-toxic | 100% | 100% |
| LTDPPGVRR | 53 | 61 | Inside | 0.488 | 0.34 | Antigen | Allergen | Non-toxic | 100% | 88.89% |
| KIAGWHGPK | 262 | 270 | Inside | 0.473 | 0.36 | Antigen | Non-antigen | Non-toxic | 100% | 100% |
| AIPITVMEY | 132 | 140 | Inside | 0.388 | 0.52 | Non-antigen | Non-allergen | Non-toxic | 100% | 100% |
| SIQDAATPY | 323 | 331 | Inside | 0.354 | 0.59 | Antigen | Non-allergen | Non-toxic | 100% | 55.56% |
| YALVDASLK | 27 | 35 | Outside | 0.287 | 0.79 | Non-antigen | Non-allergen | Non-toxic | 100% | 77.78% |
| ITQFILEHR | 201 | 209 | Inside | 0.279 | 0.81 | Non-antigen | Allergen | Non-toxic | 100% | 100% |

**Table 08**. MHC class-II epitope prediction and topology, antigenicity, allergenicity, toxicity and conservancy analysis of the epitopes of envelope protein D.

| Envelope glycoprotein D | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Epitope | Start | End | Topology | AS | Percentile scores | Antigenicity (tumor model, threshold= 0.4) | Allergenicity | Toxicity | Conservancy | Minimum identity |
| WFRMGGNCAIPI | 124 | 135 | Inside | 3.00 | 3.30 | Antigen | Allergen | Non-toxic | 100% | 83.33% |
| FRMGGNCAIPIT | 125 | 136 | Outside | 3.00 | 3.30 | Antigen | Allergen | Non-toxic | 100% | 83.33% |
| QAYQQGVTVDSI | 231 | 242 | Inside | 2.28 | 6.50 | Antigen | Non-allergen | Non-toxic | 100% | 91.67% |
| PQAYQQGVTVDS | 230 | 241 | Inside | 2.28 | 6.50 | Antigen | Allergen | Non-toxic | 100% | 83.33% |
| WNYYDSFSAVSE | 160 | 171 | Outside | 2.10 | 7.60 | Antigen | Allergen | Non-toxic | 100% | 91.67% |
| RWNYYDSFSAVS | 159 | 170 | Inside | 2.10 | 7.60 | Antigen | Allergen | Non-toxic | 100% | 91.67% |
| YWMRRHTQKAPK | 362 | 373 | Inside | 1.80 | 9.60 | Antigen | Non-allergen | Non-toxic | 100% | 58.33% |
| WMRRHTQKAPKR | 363 | 374 | Inside | 1.80 | 9.60 | Non-antigen | Non-allergen | Non-toxic | 100% | 66.67% |
| VLLNAPSEAPQI | 94 | 105 | Outside | 1.70 | 11.00 | Antigen | Non-allergen | Non-toxic | 100% | 91.67% |
| QPELAPEDPEDS | 290 | 301 | Outside | 1.70 | 11.00 | Non-antigen | Non-allergen | Non-toxic | 100% | 91.67% |

**Table 09**. B-cell epitope prediction and antigenicity and allergenicity analysis of the epitopes of the three selected proteins.

| Envelope glycoprotein E | | | | Envelope glycoprotein B | | | | Envelope glycoprotein D | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Epitope | Antigenicity (tumor model, threshold 0.4) | Allergenicity | Topology | Epitope | Antigenicity (tumor, threshold 0.4) | Allergenicity | Topology | Epitope | Antigenicity (tumor, threshold 0.4) | Allergenicity | Topology |
| PECLSPADAPCAAST | Antigen | Non-allergen | Outside | EQPRRCPTRPEGQNYT | Non-antigen | Allergen | | GLPDPFQPPSLP | Antigen | Non-allergen | |
| VVEQPLPQRGADLAEPTHPHVGAPPHAPPTHG | Antigen | Non-allergen | Outside | TTKARATAPTTRN | Antigen | Non-allergen | | APSEAPQIVRGASEDVRKQP | Antigen | Allerggen | |
| AYAPPAPSATGGL | Non-antigen | Allergen | Outside | REQSRKPPNPTPPPPGASANAS | Antigen | Non-allergen | | SACLSPQAYQQGVT | Antigen | Non-allergen | |
| AYAPPAPSATGGL | Non-antigen | Allergen | Outside | YEDQGPLVEGQLGEN | Non-antigen | Allergen | | WHGPKAPYTSTLLPPELSETPNATQPELAPEDPEDSALLEDPVGTVAPQ | Antigen | Non-allergen | |

| | | | | | | | | PPNWHIPSIQ DAATPYHPP ATPNN | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| YADWSSDSE GERDQVPWL APPERPDSPS TNGSG | Antigen | Non-allergen | Outside | RTEHKA KKKG | Antigen | Non-allergen | | IREDDQPSSH Q | Non-antigen | Allergen | |
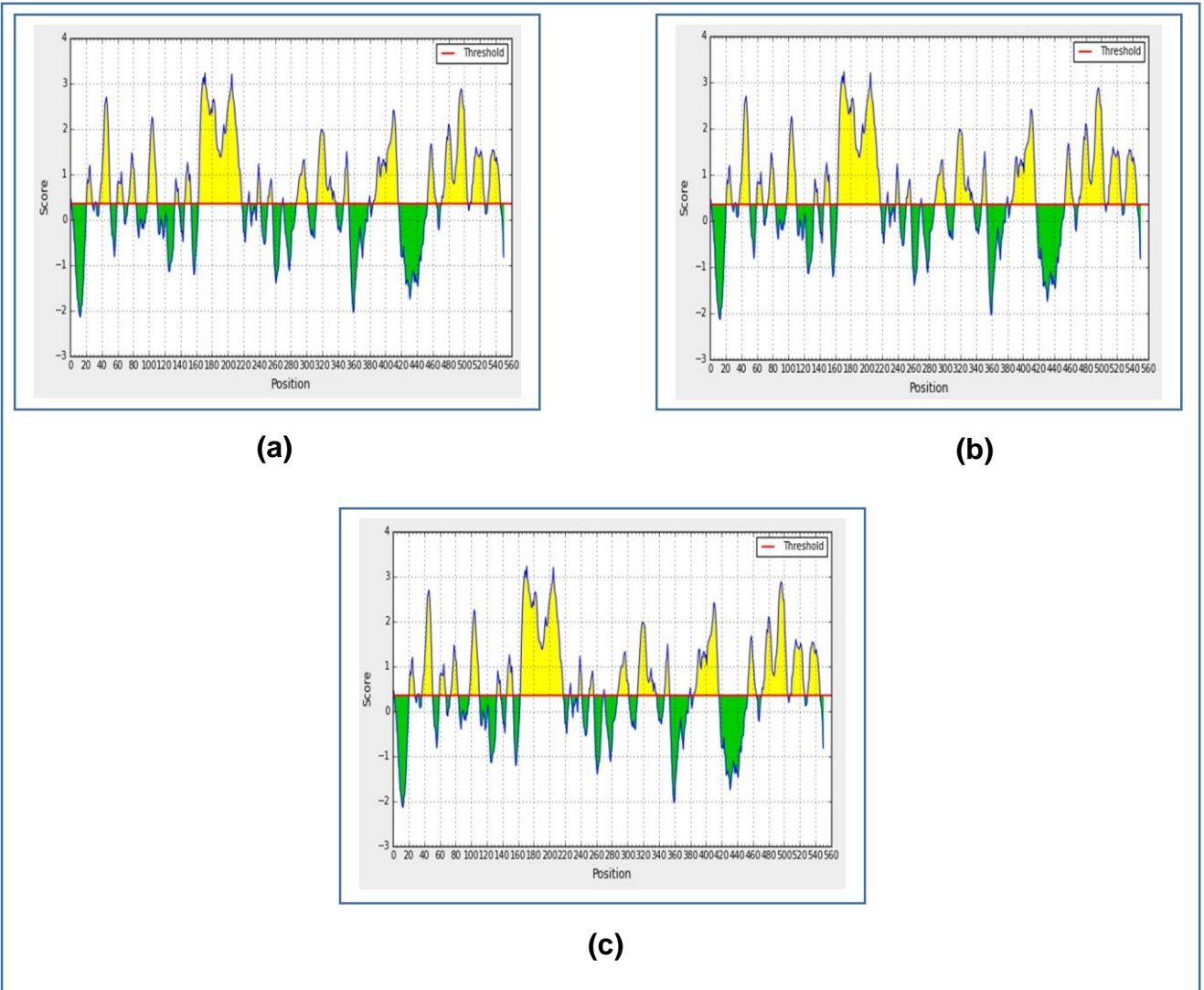
**Figure 02.** Figure showing the graphs of the B-cell epitope prediction of the three selected proteins of Human Herpes Simplex Virus-1, strain-17, using Bipipered linear epitope prediction method. Here, (a) is the graph of epitope prediction for envelope protein E, (b) is the graph of epitope prediction for envelope glycoprotein B and (c) is the graph of epitope prediction for envelope protein D.

### 3.4. Antigenicity, Allergenicity, Toxicity and Conservancy Analysis

In the antigenicity, allergenicity, toxicity and conservancy analysis, the T-cell epitopes that were found to be highly antigenic as well as non-allergenic, non-toxic, had minimum identity of over 50% and had conservancy of over 90% and the antigenic as well as non-allergenic B-cell epitopes were selected for further analysis and vaccine construction. Among the ten selected MHC class-I epitopes and ten selected MHC class-II epitopes of envelope glycoprotein E, total six epitopes (three epitopes from each of the category) were selected based on the mentioned criteria: YTLSVGDIK, ITISTAAQY, AVKSRASGK, VRMETPEAILFS, LRFDVPTSCAEM and PSWVSLMPPKQV. On the other hand, among the ten selected MHC class-I epitopes and ten selected MHC class-II epitopes of envelope glycoprotein B, total six epitopes (three epitopes from each of the category) were selected based on the mentioned criteria: KVTDMVMRK, YAYSHQLSR, ASANASVER, TMYYKDVTVSQV, YARDLTTKARAT and VMRLQSNPMKAL. Moreover, like these proteins, six epitopes that obeyed the mentioned criteria, were selected for further analysis from the envelope glycoprotein D epitopes: KMADPNRFR, KIAGWHGPK, SIQDAATPY, QAYQQGVTVDSI, YWMRRHTQKAPK and VLLNAPSEAPQI. For the selection of the B-cell epitopes, the highly antigenic and non-allergenic sequences were taken for vaccine construction. Three epitopes from each of the protein category were selected. For this reason, total nine epitopes B-cell epitopes were selected for vaccine construction, since they obeyed the selection criteria. (**Table 03**, **Table 04, Table 05, Table 06, Table 07**, **Table 08** and **Table 09**).

### 3.5. Cluster Analysis of the MHC Alleles

The cluster analysis of the possible MHC class-I and MHC class-II alleles that may interact with the predicted epitopes of the three selected proteins, were performed by online tool MHCcluster 2.0 (http://www.cbs.dtu.dk/services/MHCcluster/). The tool generates the relationship of the clusters of the alleles in phylogenetic manner. **Figure 03** illustrates the outcome of the experiment where the red zone indicates strong interaction and the yellow zone corresponds to weaker interaction.
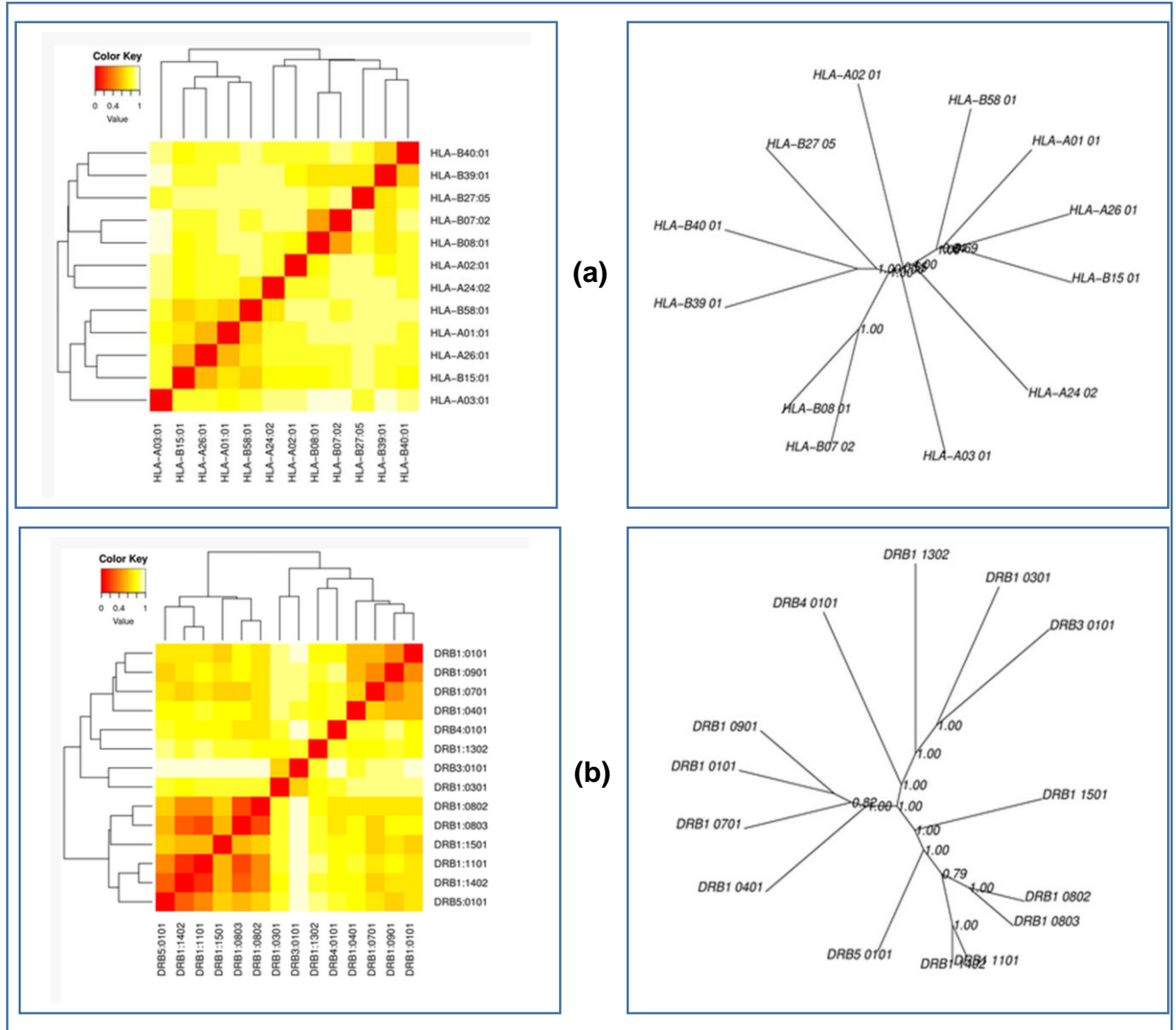
**Figure 03**. The results of the MHC cluster analysis. Here, (a) is the heat map (left) and the tree map (right) of MHC class-I cluster analysis, (b) is the heat map (left) and the tree map (right) of MHC class-II cluster analysis. The cluster analysis was carried out using onine server MHCcluster 2.0 (http://www.cbs.dtu.dk/services/MHCcluster/).

### 3.6. Generation of the 3D Structures of the Epitopes and Peptide-Protein Docking

The 3D structures of the selected T-cell epitopes were generated by the PEP-FOLD3 server. The 3D structures were generated for peptide-protein docking. The docking was carried out to find out, whether all the epitopes had the capability to bind with the MHC class-I and MHC class-II molecules. All the MHC class-I epitopes were docked against the HLA-A*11-01 allele (PDB ID: 5WJL) and MHC class-II HLA DRB1*04-01 (PDB ID: 5JLZ). The docking was performed using PatchDock server and then the results were refined by FireDock online server. Among the MHC class-I epitopes of envelope glycoprotein E, YTLSVGDIK showed the best result with the lowest global energy of -41.93. Among the MHC class-I epitopes of envelope glycoprotein B, ASANASVER generated the lowest and best global energy score of -36.46. KIAGWHGPK generated the best global energy score of -41.47 of the MHC class-I epitopes of envelope glycoprotein D. Among the MHC class-II epitopes of envelope glycoprotein E, PSWVSLMPPKQV generated the best global energy score of -7.76. TMYYKDVTVSQV generated the lowest global energy of -54.94 and QAYQQGVTVDSI generated the lowest global energy of -4.25, among the MHC class-II epitopes of envelope glycoprotein B and envelope glycoprotein D, respectively (**Table 10**).
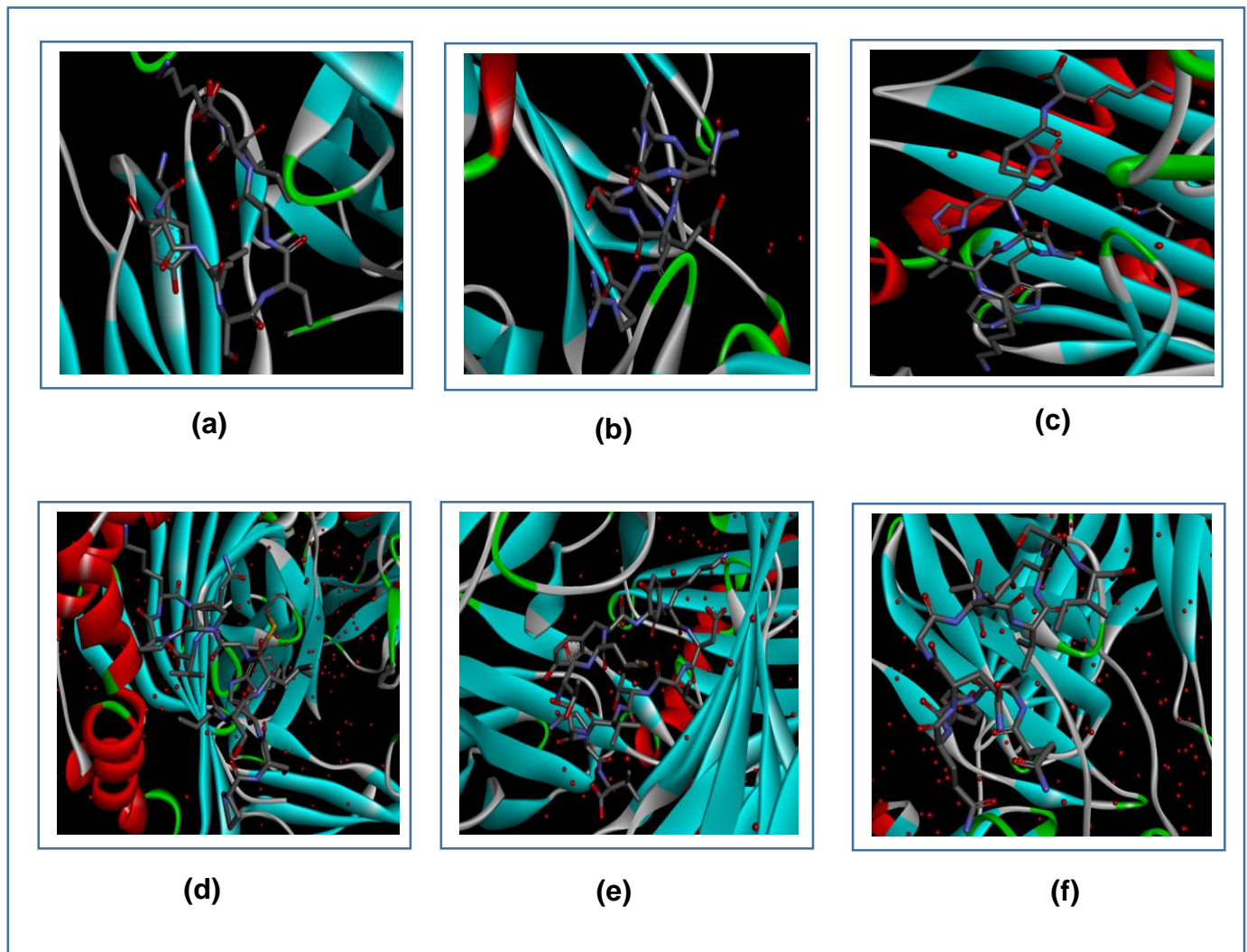
**Figure 04**. Figure showing the interactions between the best epitopes from the three proteins and their respective receptors. Here, (a) is the interaction between YTLSVGDIK and MHC class-I, (b) is the interaction between ASANASVER and MHC class-I, (c) is the interaction between KIAGWHGPK and MHC class-I molecule, (d) is the interaction between PSWVSLMPPKQV and MHC class-II, (e) is the interaction between TMYYKDVTVSQV and MHC class-II, (f) is the interaction between QAYQQGVTVDSI and MHC class-II molecule. The interactions were visualized by Discovery Studio Visualizer.

**Table 10**. Results of molecular docking analysis of the selected epitopes.

| Name of the protein | Epitope | MHC allele | Global energy | Hydrogen bond energy | Epitope | MHC allele | Global energy | Hydrogen bond energy |
|---|---|---|---|---|---|---|---|---|
| Envelope glycoprotein E | YTLSVGDIK | HLA-A*11-01 allele (PDB ID: 5WJL) | -41.93 | -3.83 | VRMETPEAILFS | HLA DRB1*04-01 (PDB ID: 5JLZ) | -3.26 | -1.00 |
| | ITISTAAQY | | -40.51 | -0.78 | LRFDVPTSCAEM | | -3.25 | 0.00 |
| | AVKSRASGK | | -17.24 | -3.26 | PSWVSLMPPKQV | | -7.76 | -0.44 |
| Envelope glycoprotein B | KVTDMVMRK | | -14.60 | -3.76 | TMYYKDVTVSQV | | -54.94 | -1.11 |
| | YAYSHQLSR | | -35.15 | -9.69 | YARDLTTKARAT | | -1.40 | 0.00 |
| | ASANASVER | | -36.46 | -2.48 | VMRLQSNPMKAL | | -5.80 | 0.00 |
| Envelope | KMADPNRFR | | -9.36 | -2.44 | QAYQQGVTVDSI | | -4.25 | -0.48 |

| glyco protei n D | KIAGW HGPK | | -41.47 | -1.89 | YWMRRHTQKAP K | | -3.13 | 0.00 |
|---|---|---|---|---|---|---|---|---|
| | SIQDAA TPY | | -26.55 | -4.32 | VLLNAPSEAPQI | | -3.43 | -0.26 |

## 3.7. Vaccine Construction

After successful docking, three vaccines were constructed, that could be used effectively to fight against HSV-1, strain-17. For vaccine constructions, three different adjuvants were used to construct three different vaccine for each of the viruses: beta defensin, L7/L12 ribosomal protein and HABA protein, were used. These three vaccines differ from each other only in their adjuvant sequences. PADRE sequence was also used for vaccine construction. Three different vaccine constructs differed from each other only in their adjuvant sequences. During vaccine construction, EAAAK, GGGS, GPGPG and KK linkers were used at their required positions. Each vaccine construct was ended by an additional GGGS linker. The newly constructed vaccines were designated as: HV-1, HV-2 and HV-3 (**Table 11**).

**Table 11**. List of the vaccines constructed against Herpes Simplex Virus-1, strain-17.

| Name of the vaccines | Vaccine constructs | Number of amino acids |
|---|---|---|
| Herpes Simplex Virus vaccine-1 (HV-1) | **EAAAK**GIINTLQKYYCRVRGGRCAVLSCLPKEEQIGKCSTRGRKCCRRKK**EAAAK**AKFVAAWTLKAAA**G****GGS**YTLSVGDIK**GGGS**ITISTAAQY**GGGS**AVKSRASGK**GGGS**KVTDMVMRK**GGGS**YAYSHQLSR**GGGS**ASANASVER**GGGS**KMADPNRFR**GGGS**KIAGWHGPK**GGGS**SIQDAATPY**GPGPG**VRMETPEAILFS**GPGP****G**LRFDVPTSCAEM**GPGPG**PSWVSLMPPKQV**GPGPG**TMYYKDVTVSQV**GPGPG**YARDLTTKARAT**GPG****PG**VMRLQSNPMKAL**GPGPG**QAYQQGVTVDSI**GPGPG**YWMRRHTQKAPK**GPGPG**VLLNAPSEAPQI**KK**P ECLSPADAPCAAST**KK**VVEQPLPQRGADLAEPTHPHVGAPPHAPPTHG**KK**YADWSSDSEGERDQVPWLA PPERPDSPSTNGSG**KK**TTKARATAPTTRN**KK**REQSRKPPNPTPPPPGASANAS**KK**RTEHKAKKKG**KK**GLP DPFQPPSLP**KK**SACLSPQAYQQGVT**KK**WHGPKAPYTSTLLPPELSETPNATQPELAPEDPEDSALLEDPVGT VAPQIPPNWHIPSIQDAATPYHPPATPNN**KK**AKFVAAWTLKAAA**GGGS** | 600 |
| Herpes Simplex Virus vaccine-2 (HV-2) | **EAAAK**MAKLSTDELLDAFKEMTLLELSDFVKKFEETFEVTAAAPVAVAAAGAAPAGAAVEAAEEQSEFD VILEAAGDKKIGVIKVVREIVSGLGLKEAKDLVDGAPKPLLEKVAKEAADEAKAKLEAAGATVTVK**EAAA****K**AKFVAAWTLKAAA**GGGS**YTLSVGDIK**GGGS**ITISTAAQY**GGGS**AVKSRASGK**GGGS**KVTDMVMRK**G****GGS**YAYSHQLSR**GGGS**ASANASVER**GGGS**KMADPNRFR**GGGS**KIAGWHGPK**GGGS**SIQDAATPY**GPGP****G**VRMETPEAILFS**GPGPG**LRFDVPTSCAEM**GPGPG**PSWVSLMPPKQV**GPGPG**TMYYKDVTVSQV**GPGP****G**YARDLTTKARAT**GPGPG**VMRLQSNPMKAL**GPGPG**QAYQQGVTVDSI**GPGPG**YWMRRHTQKAPK**GP****GPG**VLLNAPSEAPQI**KK**PECLSPADAPCAAST**KK**VVEQPLPQRGADLAEPTHPHVGAPPHAPPTHG**KK**YA DWSSDSEGERDQVPWLAPPERPDSPSTNGSG**KK**TTKARATAPTTRN**KK**REQSRKPPNPTPPPPGASANAS**K****K**RTEHKAKKKG**KK**GLPDPFQPPSLP**KK**SACLSPQAYQQGVT**KK**WHGPKAPYTSTLLPPELSETPNATQPE LAPEDPEDSALLEDPVGTVAPQIPPNWHIPSIQDAATPYHPPATPNN**KK**AKFVAAWTLKAAA**GGGS** | 685 |
| Herpes Simplex Virus vaccine-3 (HV-3) | **EAAAK**MAENPNIDDLPAPLLAALGAADLALATVNDLIANLRERAEETRAETRTRVEERRARLTKFQEDLP EQFIELRDKFTTEELRKAAEGYLEAATNRYNELVERGEAALQRLRSQTAFEDASARAEGYVDQAVELTQE ALGTVASQTRAVGERAAKLVGIEL**EAAAK**AKFVAAWTLKAAA**GGGS**YTLSVGDIK**GGGS**ITISTAAQY**G****GGS**AVKSRASGK**GGGS**KVTDMVMRK**GGGS**YAYSHQLSR**GGGS**ASANASVER**GGGS**KMADPNRFR**GG****GS**KIAGWHGPK**GGGS**SIQDAATPY**GPGPG**VRMETPEAILFS**GPGPG**LRFDVPTSCAEM**GPGPG**PSWVSL | 714 |

MPPKQV**GPGPG**TMYYKDVTVSQV**GPGPG**YARDLTTKARAT**GPGPG**VMRLQSNPMKAL**GPGPG**QAYQ QGVTVDSI**GPGPG**YWMRRHTQKAPK**GPGPG**VLLNAPSEAPQI**KK**PECLSPADAPCAAST**KK**VVEQPLPQ RGADLAEPTHPHVGAPPHAPPTHG**KK**YADWSSDSEGERDQVPWLAPPERPDSPSTNGSG**KK**TTKARATAP TTRN**KK**REQSRKPPNPTPPPPGASANAS**KK**RTEHKAKKKG**KK**GLPDPFQPPSLP**KK**SACLSPQAYQQGVT **KK**WHGPKAPYTSTLLPPELSETPNATQPELAPEDPEDSALLEDPVGTVAPQIPPNWHIPSIQDAATPYHPPAT PNN**KK**AKFVAAWTLKAAA**GGGS**

**3.8. Antigenicity, Allergenicity and Physicochemical Property Analysis of the Vaccine Constructs**

The three vaccines were found to be potent antigen as well as non-allergen. Since they are found to be non-allergenic, they are safe to use. In the physicochemical property analysis, the number of amino acids, molecular weight, extinction coefficient (in M-1 cm-1)**,** theoretical pI, half-life, aliphatic index and GRAVY were determined. All the vaccines quite similar theoretical pI and ext. coefficient (HV-3 had the lowest value of 74566.83 M-1 cm-1). All of the vaccine constructs had the similar half-life of 1 hour in the mammalian cells. HV-2 had the highest GRAVY value of - 0.544. The antigenicity, allergenicity and physicochemical property analysis of the three vaccine constructs are listed in **Table 12.**

**Table 12**. Antigenicity, allergenicity and physicochemical property analysis of the three vaccine constructs.

| Name of the vaccine | Antigenicity (on tumor model, threshold 0.4) | Allergenicity | Total amino acids | Molecular weight | Theoretical pI | Ext. coefficient (in $M^{-1}$ $cm^{-1}$) | Est. half-life (in mammalian cell) | Aliphatic index | Grand average of hydropathicity (GRAVY) |
|---|---|---|---|---|---|---|---|---|---|
| HV-1 | Antigen | Non-allergen | 600 | 62099.38 | 9.86 | 73965 | 1 hour | 51.68 | -0.712 |
| HV-2 | Antigen | Non-allergen | 685 | 70378.67 | 9.39 | 70610 | 1 hour | 60.55 | -0.544 |
| HV-3 | Antigen | Non-allergen | 714 | 74566.83 | 9.42 | 75080 | 1 hour | 59.48 | -0.665 |

### 3.9. Secondary and Tertiary Structure Prediction of the Vaccine Constructs

The secondary structures of the three vaccine constructs were generated by the online tools, PRISPRED (http://bioinf.cs.ucl.ac.uk/psipred/) and NetTurnP v1.0 (http://www.cbs.dtu.dk/services/NetTurnP/). From the secondary structure analysis, it was analyzed that, the HV-1 had the highest percentage of the amino acids (74.1%) in the coil formation as well as the highest percentage of amino acids (10.8%) in the beta-strand formation. However, HV-3 had the highest percentage of 30.1% of amino acids in the alpha-helix formation (**Figure 05** and **Table 13**).

The 3D structures of the vaccine constructs were predicted by the online server RaptorX (http://raptorx.uchicago.edu/). All the three vaccines had 3 domains and HV-2 had the lowest p-value of 8.91e-05. The homology modeling of the three dengue vaccine constructs were carried out using 1KJ6A (for HV-1), 1DD3A (for HV-2) and 4TQLA (for HV-3) as templates from protein data bank (https://www.rcsb.org/). The results of the 3D structure analysis are listed in **Table 14** and illustrated in **Figure 06.**

**Table 13**. Results of the secondary structure analysis of the vaccine constructs.

| Name of the vaccine | Alpha helix (percentage of amino acids) | Beta sheet (percentage of amino acids) | Coil structure (percentage of amino acids) |
|---|---|---|---|
| HV-1 | 15% | 10.8% | 74.1% |
| HV-2 | 24.0% | 9.9% | 66.0% |
| HV-3 | 30.1% | 6.4% | 63.4% |

**Table 14**. Results of the tertiary structure analysis of the vaccine constructs.

| Name of the vaccine | Number of the domains | p-value |
|---------------------|-----------------------|---------|
| HV-1 | 3 | 1.67e-04 |
| HV-2 | 3 | 8.91e-05 |
| HV-3 | 3 | 1.73e-04 |

(a)

(b)

(c)

**Figure 05**. The results of the secondary structure prediction of the constructed Herpes vaccines.

Here, (a) is HV-1, (b) is HV-2, (c) is HV-3. The secondary structures were predicted using online

server PRISPRED (http://bioinf.cs.ucl.ac.uk/psipred/).

**(a)**

**(b)**

**(c)**
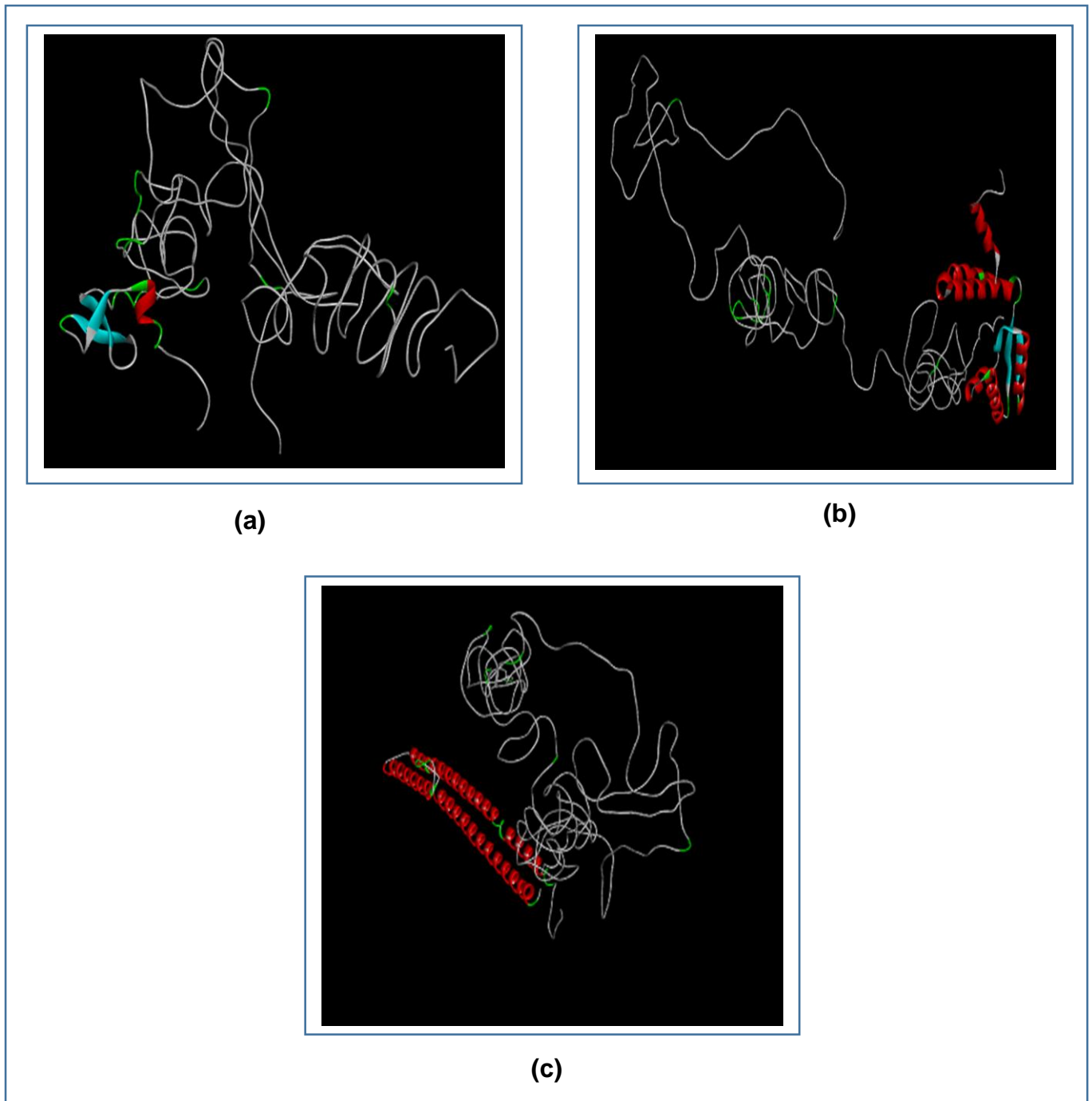
**Figure 06**. The tertiary structures of the three Herpes vaccines. Here, (a) is the HV-1, (b) is the HV-2 and (c) is the HV-3. The tertiary structures were predicted using the online server tool RaptorX (http://raptorx.uchicago.edu/) and visualized by Discovery Studio Visualizer.

### 3.10. Protein 3D Structure Refinement and Validation

The protein structures generated by the RaptorX server (http://raptorx.uchicago.edu/) were refined using 3Drefine (http://sysbio.rnet.missouri.edu/3Drefine/) and then the refined structures were analyzed by Ramachandran plot generated by PROCHECK server (https://servicesn.mbi.ucla.edu/PROCHECK/). The analysis showed that HV-1 vaccine had 63.9% of the amino acids in the most favored region, 30.6% of the amino acids in the additional allowed regions, 3.7% of the amino acids in the generously allowed regions and 1.8% of the amino acids in the disallowed regions. The HV-2 vaccine had 70.4% of the amino acids in the most favored regions, 26.2% of the amino acids in the additional allowed regions, 2.9% of the amino acids in the generously allowed regions and 0.5% of the amino acids in the disallowed regions. The HV-3 vaccine had 74.7% of the amino acids in the most favored regions, 23.7% of the amino acids in the additional allowed regions, 1.1% of the amino acids in the generously allowed regions and 0.5% of the amino acids in the disallowed regions (**Figure 07).**
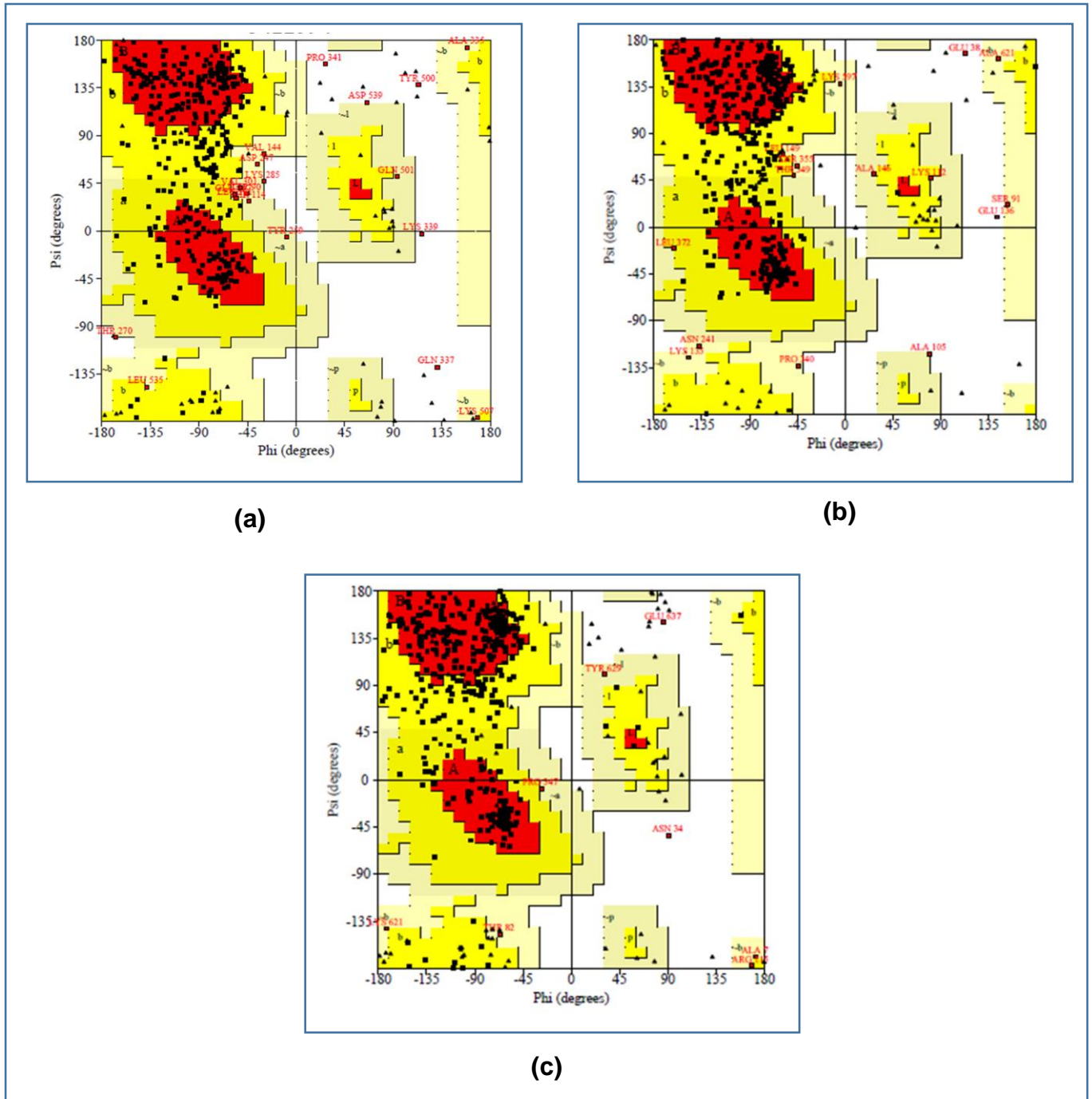
**Figure 07**.The Ramachandran plot analysis of the three vaccine constructs. (a) HV-1, (b) HV-2, (c) HV-3. The 3D structures of the constructed vaccines were refined using online refinement tool 3Drefine (http://sysbio.rnet.missouri.edu/3Drefine/) and validated by analyzing the Ramachandran plot, generated using the online tool PROCHECK (https://servicesn.mbi.ucla.edu/PROCHECK/).

### 3.11. Protein Disulfide Engineering

In protein disulfide engineering, disulfide bonds were generated for the 3D structures of the vaccine constructs. The DbD2 server identifies the pairs of amino acids that have the capability to form disulfide bonds based on the given selection criteria. In this experiment, we selected only those amino acid pairs that had bond energy value was less than 2.00 kcal/mol. The HV-1 generated 12 amino acid pairs that had the capability to form disulfide bonds. However, only 4 pairs were selected since they had bond energy, less than 2.00 kcal/mol: 19 Arg and 148 Gly, 21 Gly and 47 Arg, 177 Val and 162 Gly, 233 Pro and 251 Ser. HV-2 generated 21 pairs of amino acids that had the capability to form disulfide bonds, however, only 6 pairs were selected: 60 Glu and 67 Glu, 68 Phe and 118 Ala, 174 Ser and 217 Ser, 275 Gly and 319 Lys, 314 Ser and 325 Pro, 360 Gly and 369 Met. HV-3 generated 15 pairs of amino acids capable of forming disulfide bonds and only 3 pairs of the amino acids were selected: 191 Val and 248 Gly, 253 Ser and 290 Ser, 255 Asn and 282 Trp. The selected amino acid pairs formed the mutant version of the original vaccines in the DbD2 server (**Figure 08**).
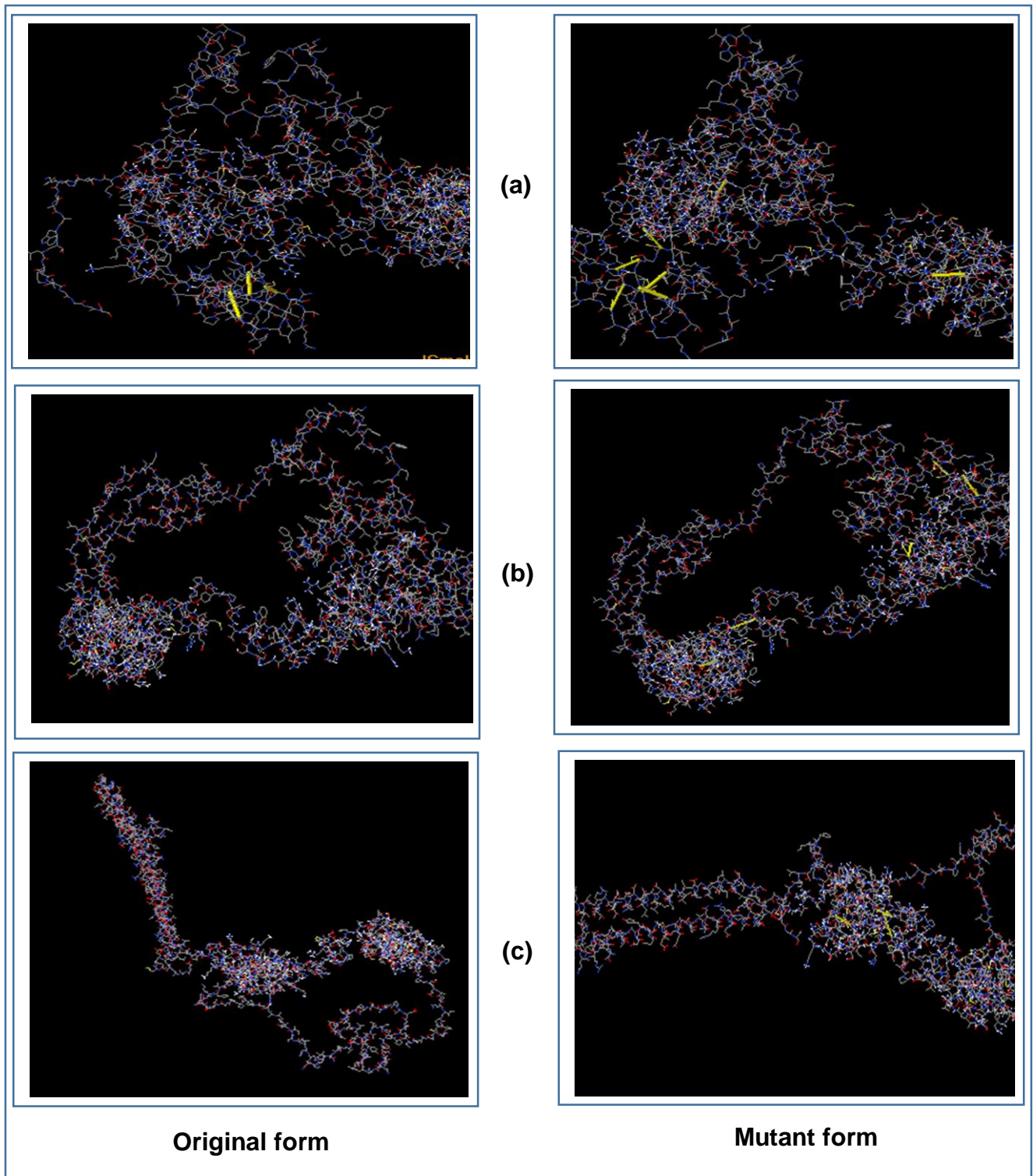
**Figure 08**. The disulfide engineering of the three vaccine constructs, both the original (left) and mutant (right) forms are shown. Here, (a) HV-1, (b) HV-2, (c) HV-3. The disulfide engineering was conducted using the online tool DbD2 server (http://cptweb.cpt.wayne.edu/DbD2/).

**3.12. Protein-Protein Docking Study**

The protein-protein docking study was carried out to find out the best constructed Herpes vaccine. From analyzing the protein-protein docking, it was declared that HV-1 was the best constructed vaccine with the best and lowest scores in the MM-GBSA study and HawkDock study. However, when analyzed by ClusPro 2.0 and later refined and re-scored by the PRODIGY tool of HADDOCK server, HV-1 showed the best binding affinity, ΔG scores with DRB3*0202 (-17.1 kcal/mol), DRB3*0101 (-19.2 kcal/mol), DRB1*0401 (-21.2 kcal/mol) and TLR-3 (-21.9 kcal/mol). And when analyzed by PatchDock and FireDock servers, HV-1 showed best global energy scores with three MHC alleles: DRB5*0101 (-18.12), DRB1*0401 (-32.33), DRB1*0301 (-13.32) and TLR-3 (-10.66). Since HV-1 showed superior results in the protein-protein docking study, it was considered as the best vaccine construct among the three constructed vaccines (**Table 15**). The molecular dynamics simulation study and in silico codon adaptation studies were conducted on only the HV-1 vaccine.

**Table 15**. Results of the docking study of all the vaccine constructs.

| Name of the vaccines | Name of the Targets | PDB IDs of the targets | Binding affinity, ΔG (kcal mol$^{-1}$) | Global energy | HawkDock score (the lowest score) | MM-GBSA (binding free energy, in kcal mol$^{-1}$) |
|---|---|---|---|---|---|---|
| HV-1 | DRB3*0202 | 1A6A | -17.1 | -7.95 | -5736.14 | -64.19 |
|  | DRB5*0101 | 1H15 | -19.5 | -18.12 | -6147.60 | -99.15 |
|  | DRB1*0101 | 2FSE | -19.3 | -8.71 | -6450.66 | -136.25 |
|  | DRB3*0101 | 2Q6W | -19.2 | 2.71 | -5975.92 | -64.87 |
|  | DRB1*0401 | 2SEB | -21.8 | -32.33 | -5086.44 | -54.85 |
|  | DRB1*0301 | 3C5J | -17.6 | -13.32 | -5850.40 | -75.96 |
|  | TLR3 | 2A0Z | -21.9 | -10.66 | -5814.57 | -81.98 |
| HV-2 | DRB3*0202 | 1A6A | -17.0 | -15.43 | -3248.84 | -38.51 |
|  | DRB5*0101 | 1H15 | -19.8 | -16.52 | -3766.88 | -32.33 |
|  | DRB1*0101 | 2FSE | -19.4 | 1.28 | -3804.33 | -16.05 |
|  | DRB3*0101 | 2Q6W | -19.0 | 7.28 | -3795.29 | -40.44 |
|  | DRB1*0401 | 2SEB | -21.3 | -5.68 | -3884.87 | -33.34 |
|  | DRB1*0301 | 3C5J | -17.5 | -7.39 | -3538.55 | -50.96 |
|  | TLR3 | 2A0Z | -20.1 | -0.00 | -3978.45 | -43.67 |
| HV-3 | DRB3*0202 | 1A6A | -16.9 | -19.01 | -3022.91 | -23.72 |
|  | DRB5*0101 | 1H15 | -19.8 | 5.56 | -3146.50 | -10.41 |
|  | DRB1*0101 | 2FSE | -19.3 | -17.60 | -2880.47 | -12.88 |
|  | DRB3*0101 | 2Q6W | -19.2 | -11.16 | 2946.17 | -6.85 |
|  | DRB1*0401 | 2SEB | -21.3 | -18.40 | -2989.85 | -13.90 |
|  | DRB1*0301 | 3C5J | -18.4 | 6.38 | -2601.94 | -19.75 |
|  | TLR3 | 2A0Z | -21.8 | -3.38 | -4212.55 | -21.56 |
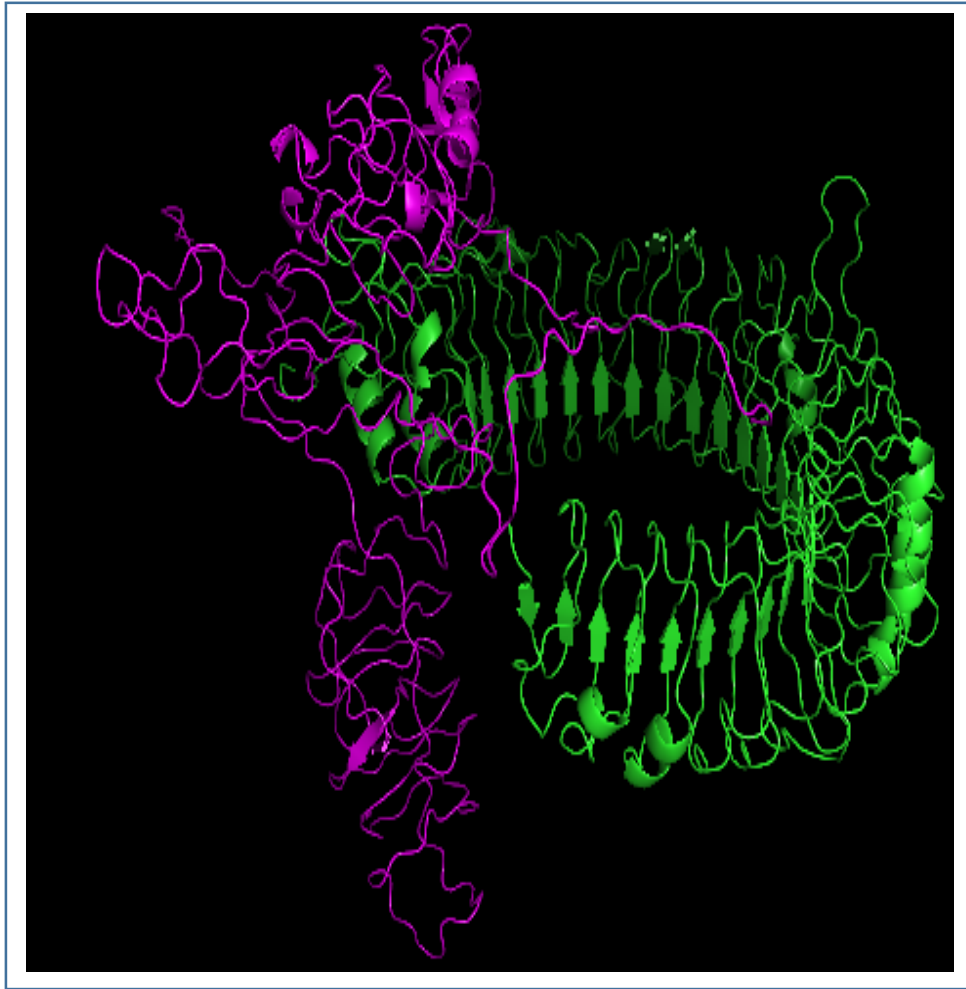
**Figure 09**. Figure showing the interaction between the TLR-3 and the selected vaccine HV-1. The interaction was visualized by PyMol tool. Here, the ligands is indicated by pink color and the receptor is indicated by green color.

### 3.13. Molecular Dynamics Simulation

**Figure 10** illustrates the results of molecular dynamics simulation and normal mode analysis (NMA) of HV-1-TLR-3 docked complex. The deformability graph of the complex indicates the peaks in the graphs which corresponds to the regions of the protein with possible deformability (**Figure 10b**). The B-factor graph of the complex gives easy visualization of the difference as well as comparison between the NMA and the PDB field of the docked complex (**Figure 10c**). The eigenvalue of the complex is illustrated in **Figure 10d**. HV-1 and TLR-3 docked complex generated eigenvalue of 1.042621e-04. The variance graph indicates the individual variance by red colored bars and cumulative variance by green colored bars (**Figure 10e**). **Figure 10f** illustrates the co-variance map of the complex where the correlated motion between a pair of residues is indicated by red color, uncorrelated motion is indicated by white color and anti-correlated motion is indicated by blue color. The elastic map of the complex represents the connection between the atoms and darker gray regions correspond to the stiffer regions (**Figure 10g**) [67, 68, 69].
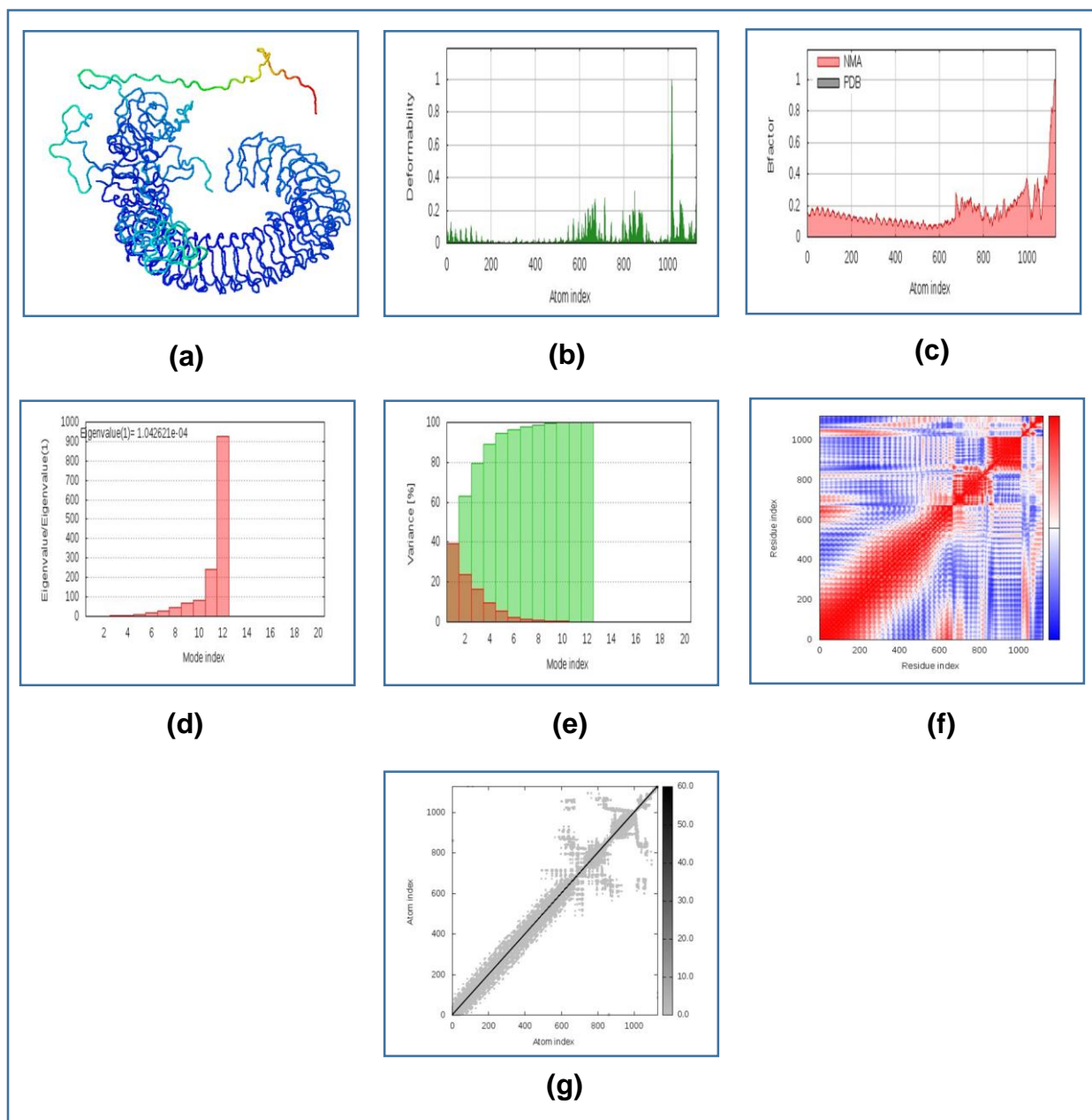
**Figure 10.** Figure displaying the results of molecular dynamics simulation study of HV-1 and TLR-3 docked complex. Here, (a) NMA mobility, (b) deformability, (c) B-factor, (d) eigenvalues, (e) variance (red color indicates individual variances and green color indicates cumulative variances), (f) co-variance map (correlated (red), uncorrelated (white) or anti-correlated (blue) motions) and (g) elastic network (darker gray regions indicate more stiffer regions).

### 3.14. Codon Adaptation and In Silico Cloning

For in silico cloning and plasmid construction, the protein sequences of the best selected vaccines were adapted by the JCat server (http://www.jcat.de/).

Since the HV-1 protein had 600 amino acids, after reverse translation, the number of nucleotides of the probable DNA sequence of HV-1 would be 1800. The codon adaptation index (CAI) value of 0.973 of HV-1 indicated that the DNA sequences contained higher proportion of the codons that are most likely to be used in the cellular machinery of the target organism *E. coli* strain K12 (codon biasness). For this reason, the production of the HV-1 vaccine would be carried out efficiently [74, 75]. The GC content of the improved sequence was 56.33%. The predicted DNA sequence of HV-1 was inserted into the pET-19b vector plasmid between the SgrAI and SphI restriction sites. Since the DNA sequence did not have restriction sites for SgrAI and SphI restriction enzymes, SgrA1 and SphI restriction sites were conjugated at the N-terminal and C-terminal sites, respectively, before inserting the sequence into the plasmid pET-19b vector. The newly constructed cloned plasmid would be 7372 base pair long, including the constructed DNA sequence of the HV-1 vaccine (the HV-1 vaccine DNA sequence also included the SgrAI and SphI restriction sites) (**Figure 11** and **Figure 12**).
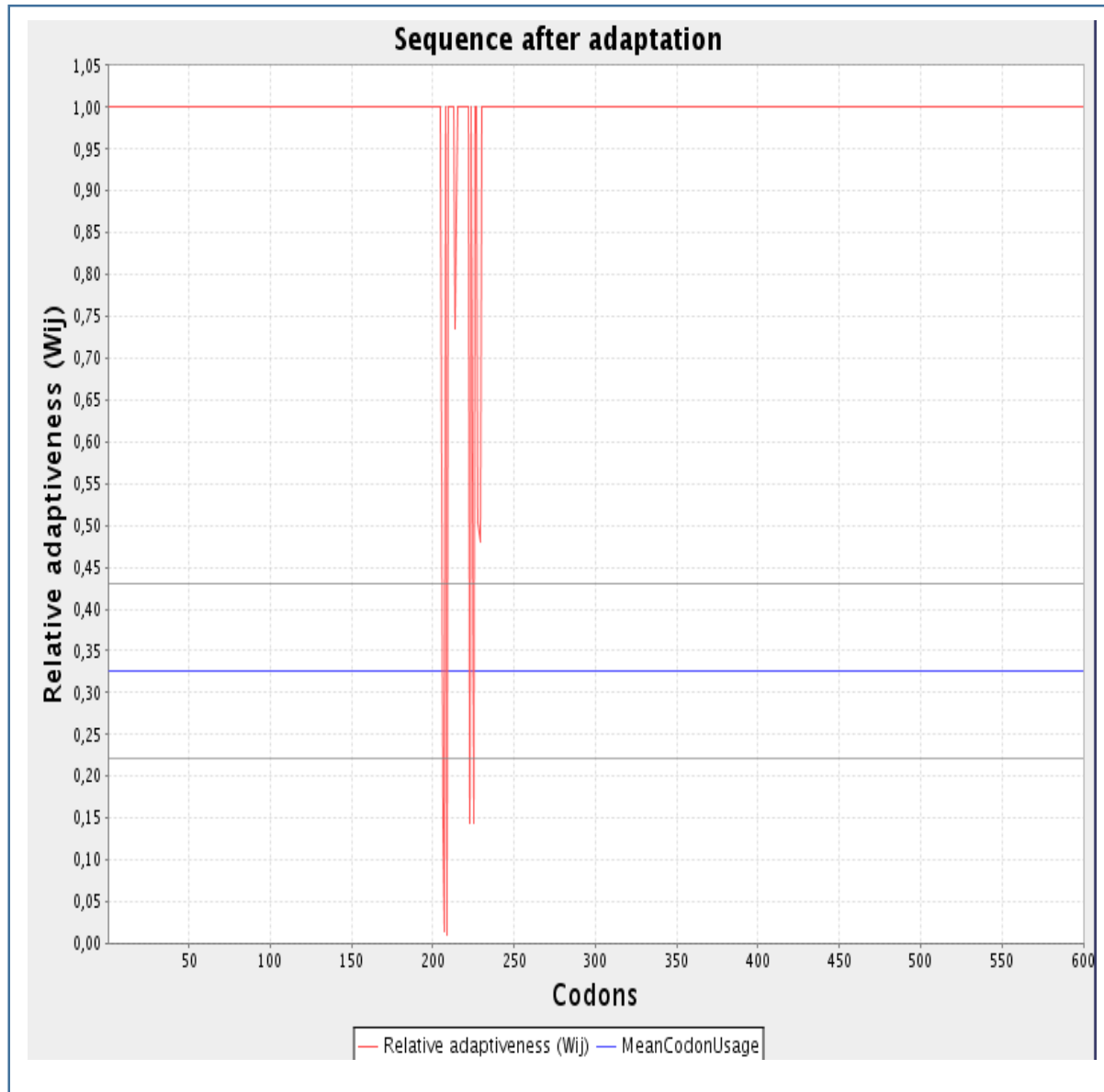
**Figure 11.** Figure showing the codon adaptation graphs of the HV-1 vaccine. The codon adaptation of the vaccine constructs were carried out using the server JCat (https://jcatbeauty.com/).
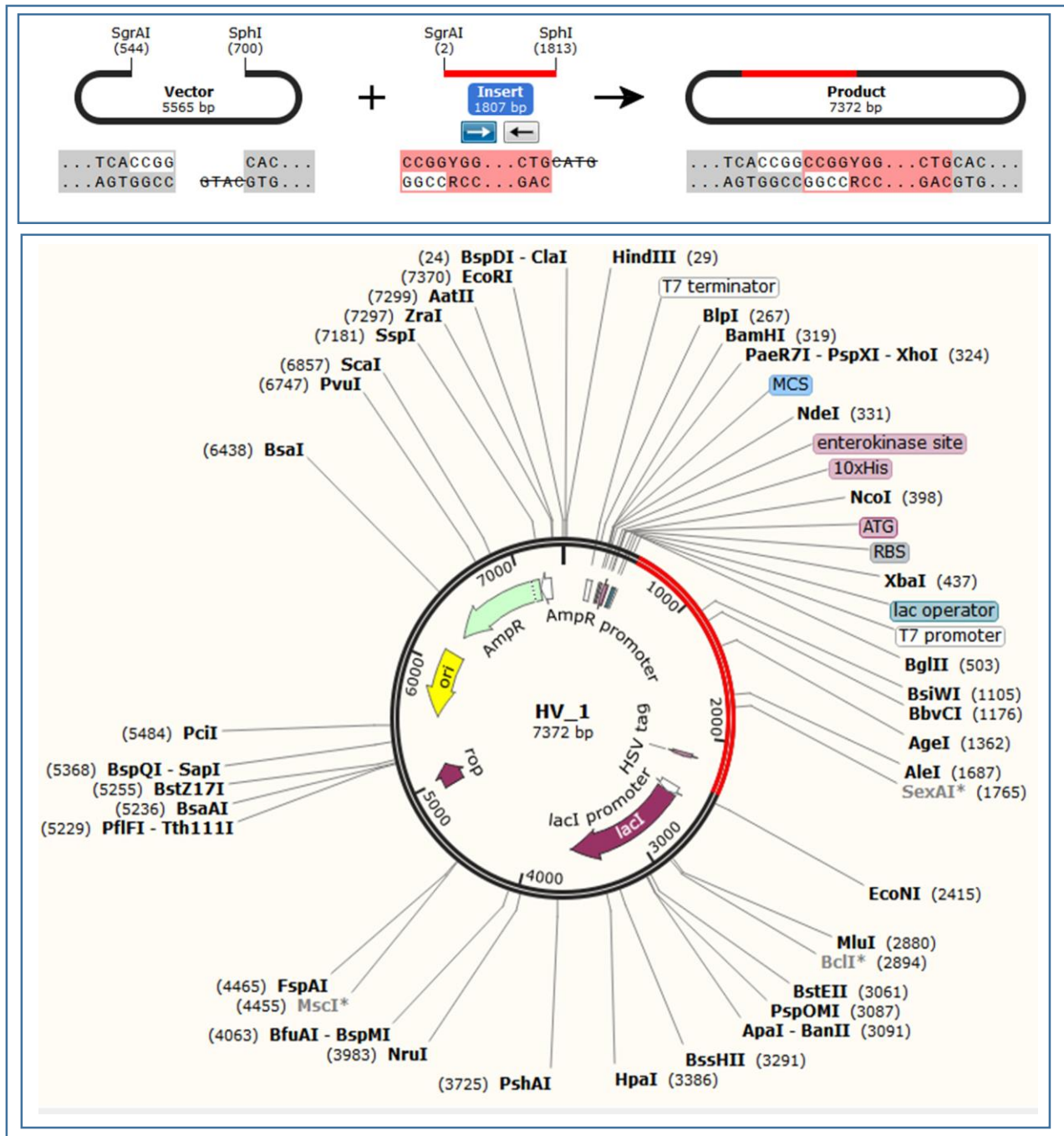
**Figure 12.** In silico restriction cloning of the HV-1 vaccine sequence in the pET-19b plasmid between the SgrAI and SphI restriction enzyme sites. The red colored marked sites contain the DNA inserts of the vaccines. The cloning was carried out using the SnapGene tool. The two newly constructed plasmids can be inserted into *E. coli* strain K12 for efficient vaccine production.

## 4. Discussion

Vaccine is one of the most important and widely produced pharmaceutical products. Millions of people and infants are getting vaccinated every year. However, the research and development processes of vaccines are costly and sometimes, it takes many years to develop a proper vaccine candidate against a particular pathogen. In recent times, various methods and tools of bioinformatics, immunoinformatics and reverse vaccinology are exploited for vaccine development, which save time and cost of the vaccine development process [76].

The current study was conducted to design and construct potential Herpes Simplex Virus (HSV), strain-17 vaccine. The target virus and its strain were identified by reviewing the NCBI database (https://www.ncbi.nlm.nih.gov/). Nine envelope glycoproteins were selected as targets for vaccine construction: envelope glycoprotein M, envelope glycoprotein H, envelope glycoprotein I, envelope glycoprotein E, envelope glycoprotein B, envelope glycoprotein L, envelope glycoprotein D, envelope glycoprotein K and envelope glycoprotein C. Envelope glycoprotein M plays important role in viral assembly [77]. The envelope glycoprotein H forms heterodimer with envelope glycoprotein L and both the envelope glycoproteins function in virus entry into the host cell. Moreover, envelope glycoprotein B, together with envelope glycoprotein H and envelope glycoprotein K, aids in the fusion events of the virion envelope with the outer nuclear membrane [78, 79, 80]. The main functions of envelope glycoprotein E and envelope glycoprotein I are to spread the virus from cell to cell as well as the fusion of the cells [81, 82, 83]. The envelope glycoprotein D also helps the virus to enter into the host cell [84]. The envelope glycoprotein C aids in the adsorption of the virion to the host cell surface [85]. Since, the glycoproteins mainly function in the infection process of the HSV, for this reason, targeting these glycoproteins could be potential strategy for vaccine construction.

After selecting the proteins, the antigenicity of the proteins were determined by VaxiJen V2.0 server. The proteins that were found to be antigenic, were selected for further analysis. Three glycoproteins, envelope glycoprotein E, envelope glycoprotein B and envelope glycoprotein D were found to be antigenic and for this reason, they were selected for further analysis. The various physicochemical properties like number of amino acids, molecular weight, theoretical pI, extinction co-efficient, instability index, aliphatic index, GRAVY were determined by ProtParam (https://web.expasy.org/protparam/) server. The three proteins performed quite similarly in the physicochemical property test.

Two of the main cells that function in immunity are the T lymphocytic cell and B lymphocytic cell. After recognized by an antigen presenting cell or APC (like macrophage, dendritic cell etc.), the antigen is presented through the MHC class-II molecule present on the surface of the APC, to the helper T cell. Since, the helper T cell contains CD4+ molecule on its surface, it is also known as CD4+ T cell. After activated by the antigen presenting cell, the T-helper cell then activates the B cell and cause the production of memory B cell and antibody producing plasma B cell. The plasma B cell produce a large number of antibodies and the memory B cell functions as the immunological memory. However, macrophage and CD8+ cytotoxic T cell are also activated by the T-helper cell, that destroy the target antigen [86, 87, 88, 89, 90]. The possible T cell and B cell epitopes of the selected viral proteins were determined by the IEDB (https://www.iedb.org/) server. The IEDB server generates and ranks the T cell epitopes based on their antigenicity scores (AS) and percentile scores. The top ten MHC class-I epitopes were taken for analysis. However, since more than one epitope showed similar AS and percentile score, two epitopes were selected from each of the AS and percentile score categories. However, based on analyzing the lengths of

the B-cell epitopes, five epitopes were selected for further analysis. The transmembrane topology of the selected eiptopes were predicted to determine whether they would be present at the exterior or interior of the viral envelope.

Antigenicity is defined as the ability of a foreign substance to act as antigen and activate the T cell and B cell responses, through their antigenic determinant portion or epitope [91]. The allergenicity of a substance is defined as the ability of that substance to act as allergen and induce potential allergic reactions within the body [92]. On the other hand, when designing a vaccine, epitopes of that remain conserved across various strains, are given much priority than genomic regions that are highly variable among the strains since the conserved epitopes of protein(s) provide broader protection across various strains and sometimes even species [93]. When selecting the best epitopes for vaccine construction, some criteria were maintained: the epitopes should be highly antigenic, so that they can induce high antigenic response, the epitopes should be non-allergenic in nature, for this reason, they would not be able to induce any allergenic reaction in an individual and the epitopes should be non- toxic. The epitopes with 100% conservancy and over 50% minimum identity were selected for vaccine construction, so that, the conserved epitopes would be able to provide protection against various strains. The conservancy analysis of the epitopes were carried out using Human Herpes Simplex Virus-1 strain-F and Human Herpes Simplex Virus-2 strain-HG52, for comparison, so that the conservancy analysis across the strains as well as species could be carried out efficiently. Total 18 T-cell epitopes (six epitopes from each of the proteins) and 9 B-cell epitopes were selected for vaccine construction. Moreover, the cluster analysis of the MHC class-I alleles and MHC class-II alleles were also carried out to determine their relationship with each other and cluster them functionally based on their predicted binding specificity [94].

In the next step, the protein-peptide docking was carried out between the selected epitopes and the MHC alleles. The MHC class-I epitopes were docked with the MHC class-I molecule (PDB ID: 5WJL) and the MHC class-II epitopes were docked with the MHC class-II molecule (PDB ID: 5JLZ). The protein-peptide docking was performed to determine the ability of the epitopes to bind with their respective MHC molecule. After 3D structure generation of the epitopes, the docking was carried out by PatchDock and FireDock servers. YTLSVGDIK, ASANASVER, KIAGWHGPK, PSWVSLMPPKQV, TMYYKDVTVSQV and QAYQQGVTVDSI generated the best scores in the protein-peptide docking. However, all of the selected T-cell epitopes showed significant interaction capability with their respective targets.

After successful protein-peptide docking, the vaccine construction procedure was carried out. Three different vaccines were constructed, that differ from each other based on their adjuvant sequences. The three vaccine constructs were designated as: HV-1 (600 amino acids long), HV-2 (685 amino acids long) and HV-3 (714 amino acids long).

After the vaccine construction was performed, the antigenicity, allergenicity and physicochemical property analysis were carried out. All the vaccine constructs were proved to be antigenic, as well as non-allergenic, for this reason, they should not cause any allergenic reaction within the body, however, all of them should be able to induce high immunogenic response. The extinction coefficient corresponds to the amount of light, that is absorbed by a compound at a certain wavelength [95, 96]. HV-3 had the highest extinction co-efficient of 74566.83 M-1 cm-1. The aliphatic index of a protein referes to the relative volume occupied by the aliphatic amino acids in

the side chains like alanine, valine etc. [97]. HV-2 had the highest aliphatic index among the vaccine constructs, although HV-3 had aliphatic index that is very near to the aliphatic index of HV-2. All the three vaccine constructs had the similar half-life of 1 hour in the mammalian cell culture and all of them were stable. All of the constructs had quite similar theoretical pI. Quite similar performances were observed by the three vaccine constructs in the physicochemical property analysis.

Two online servers, PRISPRED (http://bioinf.cs.ucl.ac.uk/psipred/) and NetTurnP v1.0 (http://www.cbs.dtu.dk/services/NetTurnP/), were used for protein secondary structure determination. HV-3 had the highest percentage of amino acids in the alpha formation and lowest percentage of amino acids in the beta-sheet formation as well as in coil structure formation. The 3D structures were generated by the server RaptorX (http://raptorx.uchicago.edu/) server and with the lowest p-value of 8.91e-05, HV-2 showed the best performance on the protein 3D structure generation. However, all of the proteins had 3 domains. After 3D structure generation, the 3D structures were refined by online tool 3Drefine (http://sysbio.rnet.missouri.edu/3Drefine/) and validated by PROCHECK (https://servicesn.mbi.ucla.edu/PROCHECK/) server. HV-3 had the best performance in the 3D structure refinement and validation study with 74.7% of the amino acids in the most favored regions and 23.7% of the amino acids in the additional allowed regions. After validation of the 3D protein structures, the disulfide engineering of the vaccine constructs were performed using Disulfide by Design 2 v12.2 (http://cptweb.cpt.wayne.edu/DbD2/) server and the amino acid pairs with binding energy value less than 2.0 kcal/mol, were chosen for disulfide bond formation. With six pairs of amino acids, that had binding energy less than 2.0

kcal/mol, it can be considered that, HV-2 showed the best performance in the protein disulfide engineering study.

In the next step, the protein-protein docking was carried out using the vaccine constructs as ligands and various MHC alleles as receptors. The docking experiment was carried out to determine the best vaccine construct among the three vaccines. All the vaccines successfully bound with their target receptors. However, from the docking experiment, it was revealed that, HV-1 was the best vaccine among the three vaccine constructs. For this reason, HV-1 vaccine construct was used for molecular dynamics simulation and in silico cloning. The molecular dynamics simulation study showed that, the HV-1-TLR-3 complex had very low chance of deformability. This was further confirmed by the deformability graph of 10b where the spikes in the graphs rarely reaches to the maximum value. The eigenvalue of the complex of 1.042621e-04 also pointed to the fact that the HV-1-TLR-3 docked complex should be quite stable and should have relatively less chance of deformability. The **Figure 10f** illustrates the correlated motion of the amino acid residue pairs in red color and according to the figure, the complex showed a large amount of amino acid pairs in the correlated motion. Finally the HV-1 vaccine was reverse transcribed into the possible DNA sequence for cloning into the *E. coli* strain K12. Before in silico cloning procedure was performed, codon adaptation of the HV-1 vaccine was carried out, where the CAI value of 0.973 was achieved, which reflects that the newly adapted sequence had high degree of codons that are mostly likely to be used by *E. coli* strain K12. Moreover, good GC content of 56.33% was also achieved. Finally, the adapted DNA sequence was inserted into the pET-19b plasmid and the plasmid with the HV-1 insert was now 7372 base pairs long. This newly generated plasmid with the target DNA sequence can be inserted into the *E. coli* strain K12 for efficient production of the HV-1 vaccine.

However, more in vitro and in vivo researches should done to finally confirm the safety, efficacy and potentiality of the predicted vaccines.

## 5. Conclusion

Herpes Simplex Virus is one of the most infectious and sexually transmitted diseases in the world. However, only the HSV-2 is transmitted sexually. The other type of Herpes virus, HSV-1, transmits though oral-oral contact. With very high infection rate, HSV can be considered as one of the fatal viruses in the world. However, no vaccines are yet discovered with satisfactory results, to control the infection of HSV. In this study, a possible subunit vaccine to fight against HSV-1, strain-17, was designed using various tools of bioinformatics, immunoinformatics and vaccinomics. In this study, first the potential proteins of the viral structure were first identified. Next, the potential epitopes were identified through robust processes and these epitopes were used for vaccine construction. Since only the highly antigenic and at the same time, non-allergenic epitopes were selected, the constructed vaccines should confer very strong antigenic response, however, with no allergenicity, the constructed vaccines should be safe to be administered. Three possible vaccines were constructed and by conducting the docking experiments, one best vaccine construct was identified. Later, the molecular dynamics simulation study and codon adaptation as well as in silico cloning were performed for large scale production of the vaccines. The vaccine should provide better immunity towards the HSV-1, strain-17 as well as the HSV-1 strain-F and the HSV-2, strain-HG52, due to its conservancy across the strains and species. However, wet lab researches should be carried out on the findings of this experiment to finally confirm the safety, efficacy and potentiality of the vaccine constructs. Hopefully, this research will raise interests among the scientists of the respective field.

**6. Acknowledgements**

Authors are thankful to Swift Integrity Computational Lab, Dhaka, Bangladesh, a virtual platform of young researchers, for providing the tools.

**7. Conflict of interest**

Bishajit Sarkar declares that he has no conflict of interest. Md. Asad Ullah declares that he has no conflict of interest.

## 8. References

[1] Whitley RJ, Roizman B. Herpes simplex virus infections. The lancet 2001;357:1513-8.

[2] Corey L, Spear PG. Infections with herpes simplex viruses. New England Journal of Medicine. 1986 Mar 13;314(11):686-91.

[3] Looker KJ, Magaret AS, May MT, Turner KM, Vickerman P, Gottlieb SL, Newman LM. Global and regional estimates of prevalent and incident herpes simplex virus type 1 infections in 2012. PloS one 2015;10:e0140765.

[4] Stevens JG, Cook ML. Latent herpes simplex virus in spinal ganglia of mice. Science. 1971;173:843-5.

[5] Looker KJ, Welton NJ, Sabin KM, Dalal S, Vickerman P, Turner KM, Boily MC, Gottlieb SL. Global and regional estimates of the contribution of herpes simplex virus type 2 infection to HIV incidence: a population attributable fraction analysis using published epidemiological data. The Lancet Infect Dis 2019.

[6] Itzhaki RF, Lin WR, Shang D, Wilcock GK, Faragher B, Jamieson GA. Herpes simplex virus type 1 in brain and risk of Alzheimer's disease. The Lancet 1997;349:241-4.

[7] Riediger C, Sauer P, Matevossian E, Müller MW, Büchler P, Friess H. Herpes simplex virus sepsis and acute liver failure. Clin Transplant 2009;23:37-41.

[8] McGrath N, Anderson NE, Croxson MC, Powell KF. Herpes simplex encephalitis treated with acyclovir: diagnosis and long term outcome. J Neurol Neurosurg Psychiatry 1997;63:321-6.

[9] Balfour HH. Resistance of herpes simplex to acyclovir. Ann Intern Med 1983;98(3):404-6.

[10] Gilbert S, McBurney E. Use of valacyclovir for herpes simplex virus-1 (HSV-1) prophylaxis after facial resurfacing: a randomized clinical trial of dosing regimens. Dermatol Surg 2000;26(1):50-4.

[11] Stanberry LR. Herpes: vaccines for HSV. Dermatol Clin 1998;16:811-6.

[12] Bernstein DI, Stanberry LR. Herpes simplex virus vaccines. Vaccine 1999;17:1681-9.

[13] Rappuoli R. Reverse vaccinology. Curr Opin Microbiol 2000;3:445-50.

[14] Chong, L.C. and Khan, A.M., 2019. Vaccine Target Discovery.

[15] Doytchinova IA, Flower DR. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. BMC Bioinformatics 2007a;8:4. doi.org/10.1186/1471-2105-8-4

[16] Doytchinova IA, Flower DR. Identifying candidate subunit vaccines using an alignment-independent method based on principal amino acid properties. Vaccine 2007b;25:856-66. doi.org/10.1016/j.vaccine.2006.09.032

[17] Doytchinova IA, Flower DR. Bioinformatic approach for identifying parasite and fungal candidate subunit vaccines. Open Vaccine J 2008;1:4.

[18] Gasteiger E, Hoogland C, Gattiker A, Wilkins MR, Appel RD, Bairoch A. Protein identification and analysis tools on the ExPASy server. InThe proteomics protocols handbook 2005;571-607. Humana press. doi.org/10.1385/1-59259-890-0:571

[19] Vita R, Mahajan S, Overton JA, Dhanda SK, Martini S, Cantrell JR, Wheeler DK, Sette A, Peters B. The immune epitope database (IEDB): 2018 update. Nucleic Acids Res 2018;47:D339-43. doi.org/10.1093/nar/gky1006

[20] Möller S, Croning MD, Apweiler R. Evaluation of methods for the prediction of membrane spanning regions. Bioinformatics 2001;17:646-53. doi.org/10.1093/bioinformatics/17.7.646

[21] Dimitrov I, Bangov I, Flower DR, Doytchinova I. AllerTOP v. 2—a server for in silico prediction of allergens. J Mol Model 2014;20:2278. doi.org/10.1007/s00894-014-2278-5

[22] Dimitrov I, Naneva L, Doytchinova I, Bangov I. AllergenFP: allergenicity prediction by descriptor fingerprints. Bioinformatics 2013;30:846-51. doi.org/10.1093/bioinformatics/btt619

[23] Thomsen M, Lundegaard C, Buus S, Lund O, Nielsen M. MHCcluster, a method for functional clustering of MHC molecules. Immunogenetics 2013;65:655-65. doi.org/10.1007/s00251-013-0714-9

[24] Lamiable A, Thévenet P, Rey J, Vavrusa M, Derreumaux P, Tufféry P. PEP-FOLD3: faster de novo structure prediction for linear peptides in solution and in complex. Nucleic Acids Res 2016;44:W449-54. doi.org/10.1093/nar/gkw329

[25] Shen Y, Maupetit J, Derreumaux P, Tufféry P. Improved PEP-FOLD approach for peptide and miniprotein structure prediction. J Chem Theory Comput 2014;10:4745-58. doi.org/10.1021/ct500592m

[26] Thevenet P, Shen Y, Maupetit J, Guyon F, Derreumaux P, Tuffery P. PEP-FOLD: an updated de novo structure prediction server for both linear and disulfide bonded cyclic peptides. Nucleic Acids Res 2012;40:W288-93. doi.org/10.1093/nar/gks419

[27] Duhovny D, Nussinov R, Wolfson HJ. Efficient unbound docking of rigid molecules. InInternational workshop on algorithms in bioinformatics 2002 Sep 17 (pp. 185-200). Springer, Berlin, Heidelberg. doi.org/10.1007/3-540-45784-4_14

[28] Schneidman-Duhovny D, Inbar Y, Nussinov R, Wolfson HJ. PatchDock and SymmDock: servers for rigid and symmetric docking. Nucleic Acids Res 2005;33:W363-7. doi.org/10.1093/nar/gki481

[29] Gujjula KR. *Prediction and comparison of hiv-1 protease inhibitor binding energies by various molecular docking methods* (Doctoral dissertation).

[30] Atapour A, Mokarram P, MostafaviPour Z, Hosseini SY, Ghasemi Y, Mohammadi S, Nezafat N. Designing a fusion protein vaccine against HCV: an in silico approach. Int J Pept Res Ther 2019;25:861-72. doi.org/10.1007/s10989-018-9735-4

[31] Dassault Systèmes BIOVIA, Discovery Studio Visualizer, v19.1.0.1828 (2019), San Diego: Dassault Systèmes.

[32] Rana A, Akhter Y. A multi-subunit based, thermodynamically stable model vaccine using combined immunoinformatics and protein structure based approach. Immunobiology 2016;221:544-57. doi.org/10.1016/j.imbio.2015.12.004

[33] Hancock RE, Nijnik A, Philpott DJ. Modulating immunity as a therapy for bacterial infections. Nat. Rev. Microbiol 2012;10:243. doi.org/10.1038/nrmicro2745

[34] Toussi DN, Massari P. Immune adjuvant effect of molecularly-defined toll-like receptor ligands. Vaccines 2014;2:323-53. doi.org/10.3390/vaccines2020323

[35] Funderburg N, Lederman MM, Feng Z, Drage MG, Jadlowsky J, Harding CV, Weinberg A, Sieg SF. Human β-defensin-3 activates professional antigen-presenting cells via Toll-like receptors 1 and 2. Proc Natl Acad Sci 2007;104:18631-5. doi.org/10.1073/pnas.0702130104

[36] Lee SJ, Shin SJ, Lee MH, Lee MG, Kang TH, Park WS, Soh BY, Park JH, Shin YK, Kim HW, Yun CH. A potential protein adjuvant derived from Mycobacterium tuberculosis Rv0652 enhances dendritic cells-based tumor immunotherapy. PloS one 2014;9:e104351. doi.org/10.1371/journal.pone.0104351

[37] Hajighahramani N, Nezafat N, Eslami M, Negahdaripour M, Rahmatabadi SS, Ghasemi Y. Immunoinformatics analysis and in silico designing of a novel multi-epitope peptide vaccine against Staphylococcus aureus. Infect Genet Evol 2017;48:83-94. doi.org/10.1016/j.meegid.2016.12.010

[38] Pandey RK, Sundar S, Prajapati VK. Differential expression of miRNA regulates T cell differentiation and plasticity during visceral leishmaniasis infection. Front Microbiol 2016;7:206. doi.org/10.3389/fmicb.2016.00206

[39] Wu CY, Monie A, Pang X, Hung CF, Wu TC. Improving therapeutic HPV peptide-based vaccine potency by enhancing CD4+ T help and dendritic cell activation. J Biomed Sci 2010;17:88. doi.org/10.1186/1423-0127-17-88

[40] Saha S, Raghava GP. AlgPred: prediction of allergenic proteins and mapping of IgE epitopes. Nucleic Acids Res 2006;34:W202-9. doi.org/10.1093/nar/gkl343

[41] Buchan DW, Jones DT. The PSIPRED protein analysis workbench: 20 years on. Nucleic Acids Res 2019;47:W402-7. doi.org/10.1093/nar/gkz297

[42] Jones DT. Protein secondary structure prediction based on position-specific scoring matrices. J Mol Biol 1999;292:195-202. doi.org/10.1006/jmbi.1999.3091

[43] Petersen B, Lundegaard C, Petersen TN. NetTurnP–neural network prediction of beta-turns by use of evolutionary information and predicted protein sequence features. PLoS One 2010;5:e15079. doi.org/10.1371/journal.pone.0015079

[44] Källberg M, Wang H, Wang S, Peng J, Wang Z, Lu H, Xu J. Template-based protein structure modeling using the RaptorX web server. Nat Protoc 2012;7:1511. doi.org/10.1038/nprot.2012.085

[45] Ma J, Wang S, Zhao F, Xu J. Protein threading using context-specific alignment potential. Bioinformatics 2013;29:i257-65. doi.org/10.1093/bioinformatics/btt210

[46] Peng J, Xu J. RaptorX: exploiting structure information for protein alignment by statistical inference. Proteins: Struct, Funct, Bioinf 2011;79:161-71. doi.org/10.1002/prot.23175

[47] Bhattacharya D, Nowotny J, Cao R, Cheng J. 3Drefine: an interactive web server for efficient protein structure refinement. Nucleic Acids Res 2016;44:W406-9. doi.org/10.1093/nar/gkw336

[48] Laskowski RA, MacArthur MW, Thornton JM. PROCHECK: validation of protein-structure coordinates. Int Tab Crystallography 2006;F:722-725.

[49] Morris AL, MacArthur MW, Hutchinson EG, Thornton JM. Stereochemical quality of protein structure coordinates. Proteins: Struct, Funct, Bioinf 1992;12:345-64. doi.org/10.1002/prot.340120407

[50] Craig DB, Dombkowski AA. Disulfide by Design 2.0: a web-based tool for disulfide engineering in proteins. BMC Bioinformatics 2013;14:346. doi.org/10.1186/1471-2105-14-346

[51] Stern LJ, Calvo-Calle JM. HLA-DR: molecular insights and vaccine design. Curr Pharm Des 2009;15:3249-61.

[52] Thompson JM, Iwasaki A. Toll-like receptors regulation of viral infection and disease. Adv. Drug Deliv Rev 2008;60:786-94. doi.org/10.1016/j.addr.2007.11.003

[53] Lester SN, Li K. Toll-like receptors in antiviral innate immunity. J Mol Biol 2014;426:1246-64. doi.org/10.1016/j.jmb.2013.11.024

[54] Mettenleiter TC. Intriguing interplay between viral proteins during herpes virus assembly or: the herpes virus assembly puzzle. Vet Microbiol 2006;113:163-9.

[55] Vajda S, Yueh C, Beglov D, Bohnuud T, Mottarella SE, Xia B, Hall DR, Kozakov D. New additions to the C lus P ro server motivated by CAPRI. Proteins: Struct, Funct, Bioinf 2017;85:435-44. doi.org/10.1002/prot.25219

[56] Kozakov D, Hall DR, Xia B, Porter KA, Padhorny D, Yueh C, Beglov D, Vajda S. The ClusPro web server for protein–protein docking. Nat Protoc 2017;12:255. doi.org/10.1038/nprot.2016.169

[57] Kozakov D, Beglov D, Bohnuud T, Mottarella SE, Xia B, Hall DR, Vajda S. How good is automated protein docking?. Proteins: Struct, Funct, Bioinf 2013;81:2159-66. doi.org/10.1002/prot.24403

[58] Vangone A, Bonvin AM. Contacts-based prediction of binding affinity in protein–protein complexes. Elife 2015;4:e07454.

[59] Xue LC, Rodrigues JP, Kastritis PL, Bonvin AM, Vangone A. PRODIGY: a web server for predicting the binding affinity of protein–protein complexes. Bioinformatics 2016;32:3676-8. doi.org/10.1093/bioinformatics/btw514

[60] Yuriev E, Ramsland PA. Latest developments in molecular docking: 2010–2011 in review. J Mo Recog 2013;26:215-39. doi.org/10.1002/jmr.2266

[61] Weng G, Wang E, Wang Z, Liu H, Zhu F, Li D, Hou T. HawkDock: a web server to predict and analyze the protein–protein complex based on computational docking and MM/GBSA. Nucleic Acids Res 2019. doi.org/10.1093/nar/gkz397

[62] Hou T, Wang J, Li Y, Wang W. Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. J Chem Inf Model 2010;51:69-82.

[63] Feng T, Chen F, Kang Y, Sun H, Liu H, Li D, Zhu F, Hou T. HawkRank: a new scoring function for protein–protein docking based on weighted energy terms. J Cheminformatics 2017;9:66. doi.org/10.1186/s13321-017-0254-7

[64] Sun H, Li Y, Tian S, Xu L, Hou T. Assessing the performance of MM/PBSA and MM/GBSA methods. 4. Accuracies of MM/PBSA and MM/GBSA methodologies evaluated by various simulation protocols using PDBbind data set. Phys Chem Chem Phys 2014;16:16719-29. DOI: 10.1039/C4CP01388C

[65] The *PyMOL* Molecular Graphics System, Version 1.2r3pre, Schrödinger, LLC.

[66] Awan F, Obaid A, Ikram A, Janjua H. Mutation-structure-function relationship based integrated strategy reveals the potential impact of deleterious missense mutations in autophagy related proteins on hepatocellular carcinoma (HCC): A comprehensive informatics approach. Int J Mol Sci 2017;18:139.  doi.org/10.3390/ijms18010139

[67] Prabhakar PK, Srivastava A, Rao KK, Balaji PV. Monomerization alters the dynamics of the lid region in Campylobacter jejuni CstII: an MD simulation study. J Biomol Struct Dyn 2016;34:778-91. doi.org/10.1080/07391102.2015.1054430

[68] López-Blanco JR, Aliaga JI, Quintana-Ortí ES, Chacón P. iMODS: internal coordinates normal mode analysis server. Nucleic Acids Res 2014;42:W271-6. doi.org/10.1093/nar/gku339

[69] Lopéz-Blanco JR, Garzón JI, Chacón P. iMod: multipurpose normal mode analysis in internal coordinates. Bioinformatics 2011;27:2843-50. doi.org/10.1093/bioinformatics/btr497

[70] Kovacs JA, Chacón P, Abagyan R. Predictions of protein flexibility: first-order measures. Proteins: Struct, Funct, Bioinf 2004;56:661-8. doi.org/10.1002/prot.20151

[71] Grote A, Hiller K, Scheer M, Münch R, Nörtemann B, Hempel DC, Jahn D. JCat: a novel tool to adapt codon usage of a target gene to its potential expression host. Nucleic Acids Res 2005;33:W526-31. doi.org/10.1093/nar/gki376

[72] Angov E. Codon usage: nature's roadmap to expression and folding of proteins. Biotechnol J 2011;6:650-9. doi.org/10.1002/biot.201000332

[73] Solanki V, Tiwari V. Subtractive proteomics to identify novel drug targets and reverse vaccinology for the development of chimeric vaccine against Acinetobacter baumannii. Sci Rep 2018;8:9044. doi.org/10.1038/s41598-018-26689-7

[74] Carbone A, Zinovyev A, Képes F. Codon adaptation index as a measure of dominating codon bias. Bioinformatics 2003;19:2005-15. doi.org/10.1093/bioinformatics/btg272

[75] Sharp PM, Li WH. The codon adaptation index-a measure of directional synonymous codon usage bias, and its potential applications. Nucleic Acids Res 1987;15:1281-95. doi.org/10.1093/nar/15.3.1281

[76] María RA, Arturo CV, Alicia JA, Paulina ML, Gerardo AO. The impact of bioinformatics on vaccine design and development. InTech, Rijeka, Croatia; 2017.

[77] Striebinger H, Funk C, Raschbichler V, Bailer S. Subcellular trafficking and functional relationship of the HSV-1 glycoproteins N and M. Viruses 2016;8(3):83. doi.org/10.3390/v8030083

[78] Farnsworth A, Wisner TW, Webb M, Roller R, Cohen G, Eisenberg R, Johnson DC. Herpes simplex virus glycoproteins gB and gH function in fusion between the virion envelope and the

outer nuclear membrane. Proceedings of the National Academy of Sciences. 2007 Jun 12;104(24):10187-92. doi.org/10.1073/pnas.0703790104

[79] Foster TP, Melancon JM, Baines JD, Kousoulas KG. The herpes simplex virus type 1 UL20 protein modulates membrane fusion events during cytoplasmic virion morphogenesis and virus-induced cell fusion. J Virol 2004;78:5347-57. doi.org/10.1128/jvi.78.10.5347-5357.2004

[80] Melancon JM, Luna RE, Foster TP, Kousoulas KG. Herpes simplex virus type 1 gK is required for gB-mediated virus-induced cell fusion, while neither gB and gK nor gB and UL20p function redundantly in virion de-envelopment. J Virol 2005;79:299-313. doi.org/10.1128/jvi.79.1.299-313.2005

[81] Han J, Chadha P, Starkey JL, Wills JW. Function of glycoprotein E of herpes simplex virus requires coordinated assembly of three tegument proteins on its cytoplasmic tail. Proceedings of the National Academy of Sciences. 2012;109(48):19798-803. doi.org/10.1073/pnas.1212900109

[82] Johnson DC, Webb M, Wisner TW, Brunetti C. Herpes simplex virus gE/gI sorts nascent virions to epithelial cell junctions, promoting virus spread. J Virol 2001;75:821-33. doi.org/10.1128/jvi.75.2.821-833.2001

[83] Johnson DC, Frame MC, Ligas MW, Cross AM, Stow ND. Herpes simplex virus immunoglobulin G Fc receptor activity depends on a complex of two viral glycoproteins, gE and gI. J Virol 1988;62:1347-54.

[84] Whitbeck JC, Peng C, Lou H, Xu R, Willis SH, De Leon MP, Peng T, Nicola AV, Montgomery RI, Warner MS, Soulika AM. Glycoprotein D of herpes simplex virus (HSV) binds directly to HVEM, a member of the tumor necrosis factor receptor superfamily and a mediator of HSV entry. J Virol 1997;71:6083-93.

[85] Herold BC, WuDUNN DA, Soltys N, Spear PG. Glycoprotein C of herpes simplex virus type 1 plays a principal role in the adsorption of virus to cells and in infectivity. J Virol 1991;65:1090-8.

[86] Goerdt S, Orfanos CE. Other functions, other genes: alternative activation of antigen-presenting cells. Immunity 1999;10:137-42. doi.org/10.1016/s1074-7613(00)80014-x

[87] Tanchot C, Rocha B. CD8 and B cell memory: same strategy, same signals. Nat Immunol 2003;4:431. doi.org/10.1038/ni0503-431

[88] Pavli P, Hume DA, Van De Pol E, Doe WF. Dendritic cells, the major antigen-presenting cells of the human colonic lamina propria. Immunology 1993;78:132.

[89] Arpin C, Dechanet J, Van Kooten C, Merville P, Grouard G, Briere F, Banchereau J, Liu YJ. Generation of memory B cells and plasma cells in vitro. Science 1995;268:720-2. DOI: 10.1126/science.7537388

[90] Cano RL, Lopera HD. Introduction to T and B lymphocytes. In Autoimmunity: From Bench to Bedside [Internet] 2013. El Rosario University Press.

[91] Fishman JM, Wiles K, Wood KJ. The acquired immune system response to biomaterials, including both naturally occurring and synthetic biomaterials. InHost Response to Biomaterials 2015;151-187. Academic Press. doi.org/10.1016/B978-0-12-800196-7.00008-6

[92] Andreae D, Nowak-Węgrzyn A. The Effect of Infant Allergen/Immunogen Exposure on Long-Term Health. InEarly Nutrition and Long-Term Health 2017;131-173. Woodhead Publishing. doi.org/10.1016/B978-0-08-100168-4.00006-9

[93] Bui HH, Sidney J, Li W, Fusseder N, Sette A. Development of an epitope conservancy analysis tool to facilitate the design of epitope-based diagnostics and vaccines. BMC Bioinformatics 2007;8:361. doi.org/10.1186/1471-2105-8-361

[94] Thomsen M, Lundegaard C, Buus S, Lund O, Nielsen M. MHCcluster, a method for functional clustering of MHC molecules. Immunogenetics 2013;65:655-65. doi.org/10.1007/s00251-013-0714-9

[95] Pace CN, Vajdos F, Fee L, Grimsley G, Gray T. How to measure and predict the molar absorption coefficient of a protein. Protein Sci 1995;4:2411-23. doi.org/10.1002/pro.5560041120

[96] Gill SC, Von Hippel PH. Calculation of protein extinction coefficients from amino acid sequence data. Anal Biochem 1989;182:319-26. doi.org/10.1093/oxfordjournals.jbchem.a133168

[97] Ikai A. Thermostability and aliphatic index of globular proteins. J Biochem 1980;88:1895-8.