Article

Immunoinformatics-Aided Analysis of RSV Fusion and Attachment Glycoproteins to Design a Potent Multi-Epitope Vaccine

Hamza Arshad Dar [1], Fahad Nasser Almajhdi [2], Shahkaar Aziz [3] and Yasir Waheed [4,*]

[1] Department of Bio and Brain Engineering, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, Korea

[2] Department of Botany and Microbiology, College of Science, King Saud University, Riyadh 11451, Saudi Arabia

[3] Institute of Biotechnology and Genetic Engineering, The University of Agriculture, Peshawar 25000, Pakistan

[4] Office of Research, Innovation & Commercialization, Shaheed Zulfiqar Ali Bhutto Medical University (SZABMU), Islamabad 44000, Pakistan

[*] Correspondence: director.oric@szabmu.edu.pk or yasir_waheed_199@hotmail.com; Tel.: +92-300-5338171

Abstract: Respiratory syncytial virus (RSV) usually causes respiratory tract infections of upper airways in infants and young children. Despite recent medical advances, no approved vaccine is available to control RSV infections. Therefore, we conducted an immunoinformatics study to design and evaluate a potential multi-epitope vaccine against RSV. Sequence-based analyses of the glycoproteins F and G revealed a total of eight CD8 T-cell and three CD4 T-cell epitopes after considering antigenicity, binding affinity and other parameters. Molecular docking analysis confirmed that these T-cell epitopes developed strong structural associations with HLA allele(s). By integrating these prioritized epitopes with linkers and a cholera toxin-derived adjuvant, a multi-epitope vaccine was designed. The developed vaccine was found to be stable, non-allergenic, flexible and antigenic. Molecular docking analysis revealed a striking mean HADDOCK score (-143.3) of top-ranked vaccine-TLR cluster and a Gibbs free energy change (ΔG) value of $-11.3\,\mathrm{kcal\,mol^{-1}}$. As per computational immune simulation results, the vaccine generated a high titer of antibodies (especially IgM) and effector T-cells. Also, codon optimization and in silico cloning ensured the increased expression of vaccine in *Escherichia coli*. Altogether, we anticipate that the multi-epitope vaccine reported in this study will stimulate humoral and cellular responses against RSV infection, subject to follow-up experimental validation.

Keywords: respiratory syncytial virus; multi-epitope vaccine; molecular modeling; molecular dynamics; vaccine candidate

1. Introduction

Human respiratory syncytial virus (RSV) belongs to the Pneumoviridae virus family [1,2]. The genome size of this species is ~15.19 kb, and it contains a single-stranded RNA with a negative sense [3,4]. A total of two subtypes of RSV have been reportedly identified, i.e., RSV-A and RSV-B [5,6]. RSV exhibits unique characteristics that distinguish it from its family members. The RSV virion is formed when the viral nucleocapsid is enclosed by a lipid envelope, that in turn is generated from the host cell's plasma membrane [7]. The RSV genome contains a total of ten genes, including those which encode fusion protein F and attachment glycoprotein G [7,8].

Using the F protein, the virus exhibits binding to the host cell surface [9]. After this successful interaction, RSV enters the host cells and forms "syncytia". The virus surface-associated fusion and glycoproteins essentially drive the synthesis of neutralizing antibodies in hosts, thus confirming their involvement in pathogenesis [7].

The virion membrane contains a metastable, prefusion form of the functional F protein trimer. Each F1 subunit has a fusion peptide (FP) at the N-terminus, which is a segment of hydrophobic residues that inserts into the target membrane. Near the C-terminus of F1, the transmembrane (TM) domain mirrors the FP, and each is linked to a heptad repeat (HR) in the following order: FP-HRA-HRB-TM. The pre-HRA refolds and trimerizes into the lengthy HRA helix when triggered [10]. The G protein is the most variable RSV protein, giving it a valuable property for research on RSV evolution. It has two large mucin-like domains that border the core area and are highly variable in their sequences. These two regions may serve as substrates for O-linked glycan decorating rather than serving any particular sequence or purpose, since the total Ser and Thr content of these two regions is rather consistent [10].

RSV is known to be very contagious [11]. The virus is transmitted primarily through respiratory droplets, direct contact with an infected individual, or exposure to surfaces contaminated with the virus [12]. RSV infections can also reoccur, as the infections weaken the long-term immunologic memory [13]. Infants are the major victims of RSV infection. While in the majority of cases, respiratory tract infections of the upper airways are caused, infections in the lower airways, such as bronchiolitis, may also occur [14]. A few symptoms are commonly observed, such as fever, sore throat, mainly nasal purulent drainage, as well as blockage of the respiratory air canals [15]. In severe cases, complications may arise, resulting in respiratory disorders such as pneumonia, asthma and bronchiolitis. The cytopathy that is caused by RSV is rather limited. It is believed that damage to air canals is brought on by immune responses, and not caused due to virus-mediated cell lysis [2].

At the time of writing, only two FDA-approved drugs were currently available to treat RSV infection [16]. In order to prevent viral infections in vulnerable children, a monoclonal antibody palivizumab is used. The use of this licensed drug has reduced hospitalization cases by 50 percent in high-risk infants [17]. Meanwhile, ribavirin has been granted FDA approval for treatment purposes in severe cases, albeit with limited effectiveness [18,19]. Due to the absence of effective treatment or prevention options to cater to all RSV infections, the development of robust therapies such as vaccines is much desired all over the world.

Research developments in the field of immunoinformatics have put immense importance on the selection of conserved epitopes (antigen-associated factors of pathogens recognized by host immune machinery) on the pathogen's antigen(s) for vaccine development [20,21]. The T-cell epitope after being presented by the relevant HLA on the target cell's surface activates the receptor present on the T-cell surface, which proliferates and ensures the induction of specific immune responses [22]. Depending on the epitope presented on the MHC surface, either a CD4+ T cell or CD8+ T cell epitope vaccine could be developed. There are many advantages associated with the use of epitope-based peptide vaccines. As they do not contain the full pathogen, their safety concerns are limited [23]. Furthermore, they are economical to produce (cheap and involve less daunting procedures). It is possible to incorporate sugar analogs in vitro in epitope vaccines. The production process takes less time compared to conventional vaccines; moreover, it results in enhanced specificity and stability [24]. The underlying principle of epitopes-based vaccines helps focus protective immune responses on designated and specified antigenic determinants, limiting the induction of unwarranted or unwanted immunological responses [25]. In immunoinformatics research, it is desirable to identify epitopes that tend to associate with a variety of HLA alleles, and thus provide protective immunity in human populations globally [26]. In this study, we conducted comprehensive immunoinformatics investigations on the RSV F and G glycoproteins in order to obtain high-ranked epitopes. The prioritized T-cell epitopes were joined together to design a multi-epitope, and then evaluated using rigorous structural modeling, molecular dynamics and immune simulations.

When microorganisms break through the mucosal barrier, Toll-like receptors (TLRs) can recognize them and initiate adaptive immune responses [27]. TLR4 has been previously investigated to recognize the viral structural proteins and stimulate the production of inflammatory cytokines [28]. Besides, several studies have indicated the importance of

TLR4 in forming a robust immune response against viral infection [29,30]. Therefore, we explored the interaction of epitope-based vaccine with TLR4 structure using molecular docking, molecular dynamics simulation and free energy-based analyses. We believe our study is a valuable addition to previous vaccine studies on RSV. Subject to experimental validation, this will be a significant step towards developing an anti-RSV vaccine.

2. Materials and Methods

2.1. Collection of the RSV Glycoproteins F and G Sequences

The reference sequence of the RSV genome (accession number NC_001803.1) was extracted from the GenBank database of NCBI. From its annotation, the amino acid sequences of both the F and G glycoproteins were collected in FASTA format.

2.2. Prioritization of T-Cell Epitopes in the RSV F and G Glycoproteins

The F and G glycoprotein sequences were fed into the NetMHC 4.0 server (http://www.cbs.dtu.dk/services/NetMHC/ accessed on 7 April 2021) in order to forecast strong CD8 T-cell binders using default settings [31,32]. HLA I alleles selected for prediction were HLA-A*1101, HLA-A*0101, HLA-A*0301, HLA-A*0201 and HLA-B*3501. Next, these epitopes were assessed for antigenicity properties as per VaxiJen 2.0 (http://www.ddg-pharmfac.net/vaxijen/ accessed on 8 April 2021) by specifying a 0.5 threshold [33]. Next, the immunogenicity scores of epitope peptides were determined using the IEDB resource [34]. Finally, the binding potentials of these epitopes with their respective allele(s) were analyzed by calculating their IC50 values, using the MHCPred 2.0 (http://www.ddg-pharmfac.net/mhcpred/ accessed on 12 April 2021) [35]. Only epitopes exhibiting strong binding potential (IC50 value < 100 nm) were prioritized. It was also confirmed that these 9-mer peptides do not show identity with any human genomic region using BLASTp against the human genome.

The F and G glycoprotein sequences were checked using the NetMHCIIPan 3.2 server (http://www.cbs.dtu.dk/services/NetMHCIIpan-3.2/ accessed on 16 April 2021) in order to identify strong CD4+ T-cell binders [36]. HLA-DRB alleles selected for this purpose were DRB1*0101, DRB1*0401 and DRB1*0101. Next, antigenicity evaluation was carried out on these epitopes by VaxiJen 2.0 by specifying a 0.5 threshold [33]. Finally, their core 9-mer peptides (identified by the NetMHCIIPan server) were analyzed for binding affinity with their putative HLA DRB allele(s) using MHCPred 2.0 (https://www.ddg-pharmfac.net/mhcpred/MHCPred/), developed by Guan, Pingping, et al., Edward Jenner Institute for Vaccine Research, Berkshire, UK (accessed on 25 April 2021) [35]. Only those epitopes showing an IC50 < 100 nm were prioritized. It was also confirmed that these 15-mer peptides do not show identity with any human genomic region using BLASTp against the human genome. Moreover, the solvent accessibility of shortlisted epitopes residues was predicted via the WESA meta-predictor (https://pipe.rcc.fsu.edu/wesa/ accessed on 6 August 2022) [37].

2.3. Molecular Docking of Epitopes with HLA Alleles

The structural modeling of finalized T-cell epitopes was performed by the PEP-FOLD 3.5 online tool, specifying default settings [38]. Only the structure showing the lowest sOPEP energy was selected. The obtained epitope structures were used to conduct molecular dockings with their putative interacting HLA allelic structures that were defined earlier. This was necessary in order to perform structural characterization of the epitope-HLA interaction at the molecular level. We used the HADDOCK 2.2 server (http://haddock.science.uu.nl/services/HADDOCK2.2/), developed by Van Zundert, et al., Utrecht University, the Netherlands (accessed on 23 May 2021). This server was used for molecular docking analysis, and in order to calculate the various statistical properties associated with the overall process [39,40]. Guru interface, an advanced interface, was utilized for this purpose with its default settings.

The structures of complete HLA with co-crystallized peptides (HLA A*0101, HLA A*0201, HLA A*0301, HLA A*1101, HLA DRA/B*0101 and HLA DRA/B*0401) were retrieved from the RCSB database. These structures were retrieved from PDB ID 4NQX, 1DUZ, 3RL1, 1Q94, 1AQD and 5JLZ, respectively [41–46]. Co-crystallized peptides containing the structure of HLA alleles were submitted to the PDBsum server to identify those residues of HLA that formed bonded and non-bonded interactions with the co-crystallized peptides [47]. These amino acid residues were fed into the HADDOCK 2.2 server as the active residues of HLA molecules to guide the process of molecular docking. In the case of HLA I 9-mer peptides, the epitope residues were fed as active residues. However, for 15-mer HLA II epitopes, their core 9-mer regions were defined as active residues, whereas their surrounding residues were automatically designated as passive. The top cluster having the minimum HADDOCK score was considered, and a representative protein-peptide complex from it was downloaded in PDB format and visualized in UCSF Chimera version 1.15, developed by Resource for Biocomputing, Visualization, and Informatics (RBVI), University of California, San Francisco, CA, USA. [48]. The active binding site residues of HLA were identified using the PDBsum analysis suite [47]. The protein-peptide complexes obtained were also visualized using the UCSF Chimera for better representation [48].

2.4. In Silico Design of RSV Multi-Epitope Vaccine

A multi-epitope vaccine was designed against RSV by incorporating all the prioritized T-cell epitopes. Using a special linker GPGPG, epitopes were joined together. Meanwhile, cholera toxin subunit B was placed towards the N-terminal end of epitopes as an adjuvant to enhance the overall immunogenicity of the vaccine. The adjuvant was connected to the first epitope by a single rigid linker EAAAK. Crucial chemical and physical properties of the designed multi-epitope were studied using the ProtParam server's Expasy tool [49]. These predictions concerned the half-life, molecular weight pH, hydropathicity, instability and aliphaticity. Furthermore, the designed poly-epitope vaccine was subjected to antigenicity checks by VaxiJen and AntigenPro servers [33,50]. In order to confirm that the vaccine does not trigger unwanted allergenic responses, AllergenFP version 1.0 and AllerTOP version 2.0 (both tools developed by Dimitrov, Ivan, et al., from Medical University of Sofia, Sofia, Bulgaria) were also used [51,52].

2.5. Structural Modeling of the RSV Multi-Epitope Vaccine

Three-dimensional modeled structures of the multi-epitope vaccine were generated by multiple servers: Phyre2, RoseTTAFold, i-TASSER, RAPTORX and 3Dpro [53–57]. All the structures that were obtained were checked for stereochemical quality and structural correctness by PROCHECK, ERRAT and ProSa-Web [58–61]. The top three modeled structures were subjected to molecular refinements by the GalaxyRefine web server (http://galaxy.seoklab.org/ accessed on 27 May 2021) [62]. Careful comparison of structures obtained before and after molecular refinements helped to select one model. In particular, Ramachandran plot analyses and ERRAT scores played a crucial role in model selection.

2.6. Molecular Docking Analysis of Vaccine with Toll-Like Receptor 4

The full-chain sequence of human TLR4 (without any signal peptides) was obtained from the UniProt sequence data in FASTA format. NCBI PSI-BLAST search was performed against this protein sequence against the Protein DataBank database. TLR4 structure with PDB ID 3FXI was downloaded from the PDB database [63]. All molecules except the TLR chain were removed using the UCSF Chimera tool [48]. The UCSF Chimera built-in option of 'Dockprep' was used with its default settings to prepare the TLR4 and multi-epitope structures before docking. For performing molecular docking, the HADDOCK 2.2 server was used [39,40]. This web server requires active and passive residues of input protein structures; these residues were identified by the CPORT online tool with a very sensitive option, as it is recommended by the HADDOCK server [64]. The HADDOCK refinement interface was used to improve the overall positioning and docking pose of the

vaccine-TLR4 docked structure [40]. Some notable interface residues within the vaccine-TLR4 complex and inter-chain bonding forces were determined by the PDBsum [47]. Thermodynamics-based calculation of the Gibb's free energy change was calculated by the PRODIGY server (http://wenmr.science.uu.nl/prodigy/ accessed on 1 June 2021), using its default parameters [65].

2.7. Molecular Dynamics Analysis on Vaccine-TLR4 Complex

GROMACS, a popular command line-based Linux tool, was used for molecular simulations on the vaccine-TLR4 complex [66]. Molecular simulations are usually recommended to imitate the environment in which biological interactions occur. This procedure also evaluates the stability of individual protein chains. Optimized Potential for Liquid simulation force field was selected [67]. The protein structural files were converted to gro format, in order to prepare the topology that is fully compatible with the chosen force field. In order to accommodate the water molecules, the vaccine-TLR4 structure was placed within the center of a cubic box having a 1-nanometer distance from the edge so that the structures' periodic images were 2 nm apart. Water molecules were position-restrained using a constant force of 1000 kJ mol^{-1} nm^{-2}. In order to neutralize the overall charge distribution in the system, the genion tool was used to add counter ions into the system. For this purpose, a Verlet cut-off scheme as well as Ewald electrostatic forces were applied.

The process of energy minimization was completed at 2098 steps. Graphs obtained during various stages of simulation and/or equilibration steps, such as energy minimization, temperature, pressure and density, were analyzed using Xmgrace tool version 5.1.19 (developed by Paul J. Turner, Center for Coastal and Land-Margin Research Oregon Graduate Institute of Science and Technology, Beaverton, OR, USA.) [68]. MD simulation was performed on the equilibrated vaccine-TLR4 structure for 20 ns. RMSD analysis concerning the backbone of the stabilized structure was carried out. During an MD production run of 20 ns, the radius of gyration parameter was also calculated to understand the compactness of the vaccine-TLR4 structure. Finally, the RMSF of side chains was also computed to check the flexibility of these regions during simulation.

2.8. In Silico Immune Simulation Analysis

In order to conduct computer-aided investigations on the immune potential of the vaccine, the C-ImmSim (http://kraken.iac.rm.cnr.it/C-IMMSIM/ accessed on 12 August 2021) server was used [69]. By combining a position-specific scoring matrix with machine learning, this server assesses epitope prediction and dynamic immunological activities. The whole analysis was performed using the protocol previously reported [70,71]. Briefly, three in silico doses were administered four weeks apart, at time steps of 1, 84, and 168 (single time step equals 8 h of daily life), respectively, for a total of 1050 simulation steps (about 12 months). Nevertheless, the default values of remaining parameters were used.

2.9. Cloning of Designed Vaccine

Using the EMBOSS Backtranseq platform, a reverse translation of the vaccine sequence was conducted [72]. Next, in order to optimize the codon sequence, the Java Codon Adaptation Tool at http://www.jcat.de/ (developed by Grote, Andreas, et al., Technische Universität Braunschweig, Germany), was utilized (accessed on 14 August 2021) [73]. This server provides Codon Adaptation Index (CAI) and GC% content, which were checked to evaluate expression levels of the vaccine. In addition, *Xho*I and *Nde*I restriction sites were added at the N-terminal and C-terminal ends of the optimized codon sequence of the vaccine candidate, respectively. Finally, the SnapGene tool was used (accessed on 16 August 2021) to clone the optimized sequence of vaccine into the expression vector pET28a (+) between *Xho*I and *Nde*I loci.

3. Results

3.1. Prioritization of T-Cell Epitopes in the RSV F and G Glycoproteins

Initially, 42 epitopes were predicted in the F and G glycoproteins that showed a strong putative association with HLA I alleles. Out of these, it was found that a total of 25 epitopes showed good VaxiJen scores (>0.5), thus indicating their antigenic tendencies. Immunogenicity evaluation of these sequences shortlisted a total of 13 epitopes. Finally, out of these, it was revealed that a total of eight epitopes had IC50 values < 100 nm, which indicates their strong binding potential with the predicted HLA I alleles (Table 1). BLASTp analyses of these peptide sequences with the human genome revealed that they do not show identity with any human genomic region.

Table 1. The overall attributes associated with the selected eight HLA I interacting T-cell epitopes. In epitope sequence, exposed residues are bold and italicized, while the rest are buried residues as predicted by the WESA tool.

Position	Epitopes	HLA Alleles	Protein	VaxiJen Score	Imunogenicity Score	IC50 Value
383–391	NIDIFNPKY	HLA-A*0101	Fusion glycoprotein	0.7826	0.11367	49.43
541–549	LIAVGLLLY		Fusion glycoprotein	0.8288	0.06096	30.13
525–533	IMITTIIIV	HLA-A*0201	Fusion glycoprotein	0.5548	0.43542	77.62
512–520	LLHNNAGK	HLA-A*0301	Fusion glycoprotein	0.5354	0.09092	60.53
132–140	KTKNTTTTK		Attachment glycoprotein	0.8053	0.05327	54.83
57–64	ITIELNIK	HLA-A*1101	Fusion glycoprotein	1.4507	0.04972	13.03
201–210	KQLPIVNK			0.8386	0.13078	27.04
148–156	IASGAVSK			0.8027	0.08501	36.48

Initially, 49 epitopes were predicted in the F and G glycoproteins that showed a strong putative association with HLA II alleles. Out of these, a total of 28 epitopes showed good VaxiJen scores (>0.5), thus indicating their antigenic tendencies. After exploring their binding affinities in terms of IC50 values of their 9-mer core regions, as well as selecting the best epitopes from among the overlapping epitopes, a total of three CD4+ T-cell epitopes were prioritized accordingly (Table 2). We also included AIIFIASANHKVTLT epitope in the DRB1*0401 results obtained, although it had an IC50 value somewhat higher with that allele, i.e., 264.24, as it had a good antigenicity score and binding affinity with another frequent HLA allele, DRB1*0101. Upon BLASTp analyses of these peptide sequences with the human genome, it was found that none showed identity with any human genomic region.

Table 2. Table showing features of the selected HLA II epitopes. In epitope sequence, exposed residues are bold and italicized, while other residues are buried (WESA tool).

Position	Epitope	HLA Alleles	Protein	VaxJen Score	IC50 Value
138–152	LGFLLGVGSAIASGI	DRB1*0101	Fusion glycoprotein	0.6271	37.33
546–560	LLLYCKARSTPVTLS	DRB1*0101		1.2472	2.75
	AIIFIASANHKVTLT	DRB1*0101	Attachment glycoprotein	0.7845	40.18
58–72	AIIFIASANHKVTLT	DRB1*0401			264.24

The selected T-cell epitopes showed good structural interactions with their cognate HLA allele(s) as indicated by their low HADDOCK scores (Supplementary Material, Table S1). Among the HLA I T-cell epitopes, the lowest HADDOCK score (−113.8) was shown by epitope KTKNTTTTK, followed by epitope LLHNVNAGK (−104.2). Figure 1 shows interactions of top-ranked HLA I epitopes with their cognate HLA allele(s). HLA II T-cell epitopes also showed good structural interactions with their cognate HLA II allele(s) (Supplementary Material, Table S1). Notably, the epitope AIIFIASANHKVTLT showed striking HADDOCK scores of −117.4 and −117.1 with HLA DRA/B1*0101 and HLA DRA/B1*0401, respectively. Best protein-epitope docked conformations can be visualized in Figure 2. These figures also visualized strong peptide-protein interactions within the binding groove of HLA molecule that is helpful for stable binding.
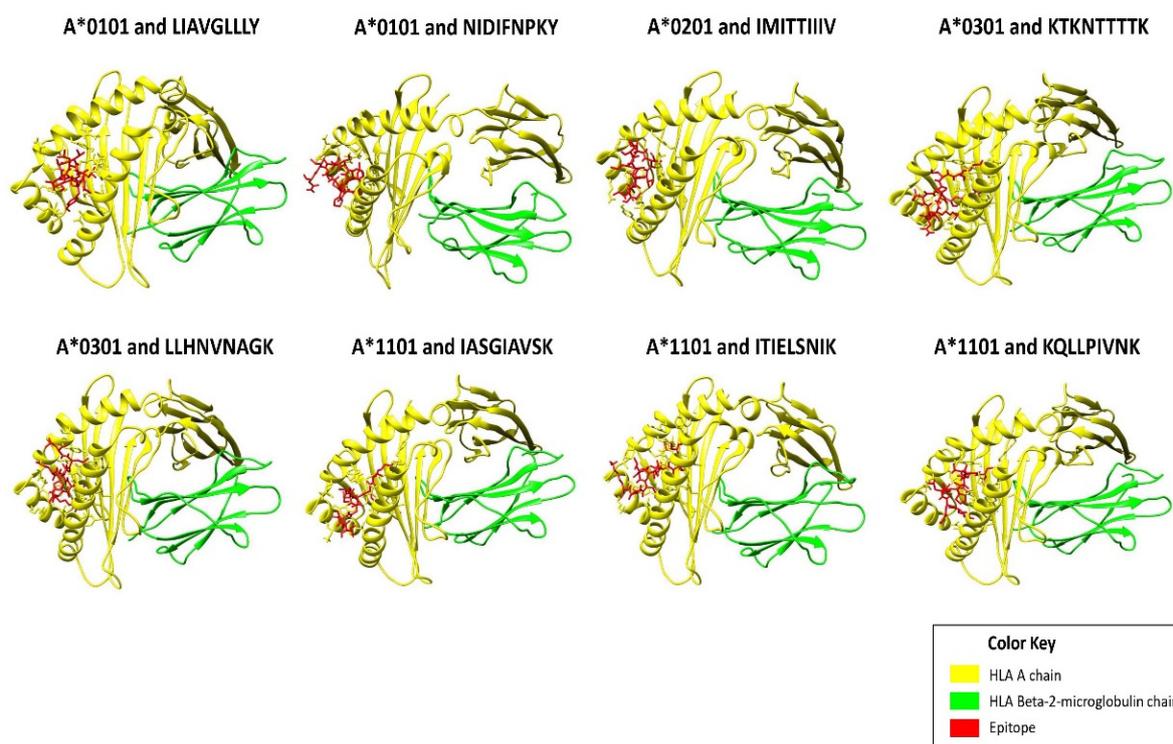
Figure 1. Figure showing structural interactions between the prioritized HLA I epitopes (colored in red) and cognate HLA I chains (colored in yellow and green). Out of all the HLA I epitopes, KTKNTTTTK (from the G glycoprotein) showed the best docking score (−113.8) with cognate HLA I molecule HLA-A*0301, followed by epitope LLHNVNAGK (−104.2) of the F glycoprotein with HLA-A*0301, and so on. It can be observed that all the epitopes fit comfortably inside the binding groove of the HLA molecule and thus are projected to interact strongly with their cognate HLA I allele.

3.2. In Silico Designing, Physiochemical and Immunological Properties Evaluation of Multi-Epitope Vaccine

The prioritized HLA II epitope LIAVGLLLYCKARSTPVTLS contained the smaller HLA I epitope LIAVGLLLY; hence, the larger HLA II epitope was retained out of these two epitopes. In addition, overlapping residues were observed between two epitopes (HLA II epitope LGFLLGVGSAIASGI and HLA I epitope IASGIAVSK), so they were merged together to form one contiguous epitope LGFLLGVGSAIASGIAVSK. By joining all the prioritized T-cell epitopes with adjuvant and linkers (GPGPG and EAAAK), a multi-epitope vaccine was designed. As per the physicochemical analysis of the vaccine, its molecular weight is 26.657 kDa. It is generally acknowledged that proteins with a lower molecular weight than 110 kDa are more likely to show antigenic capabilities [21,74]. It was predicted that the in vitro half-life of the vaccine will be 30 h in mammalian reticulocytes, the in vivo

half-life will exceed 20 h in yeast, and the in vivo half-life will exceed 10 h in *Escherichia coli*. As per server calculations, the instability index of protein is 26.61 (<40), thus stability is anticipated. A high value of 93.81 for the aliphatic index was obtained, which indicates increased thermostability at varying temperatures. A GRAVY score of 0.003 was obtained; thus, it was concluded that the vaccine is slightly prone to being hydrophobic. Therefore, the use of a stable, lipid-based reconstitution system would be helpful to overcome the protein's self-assembly [75,76].
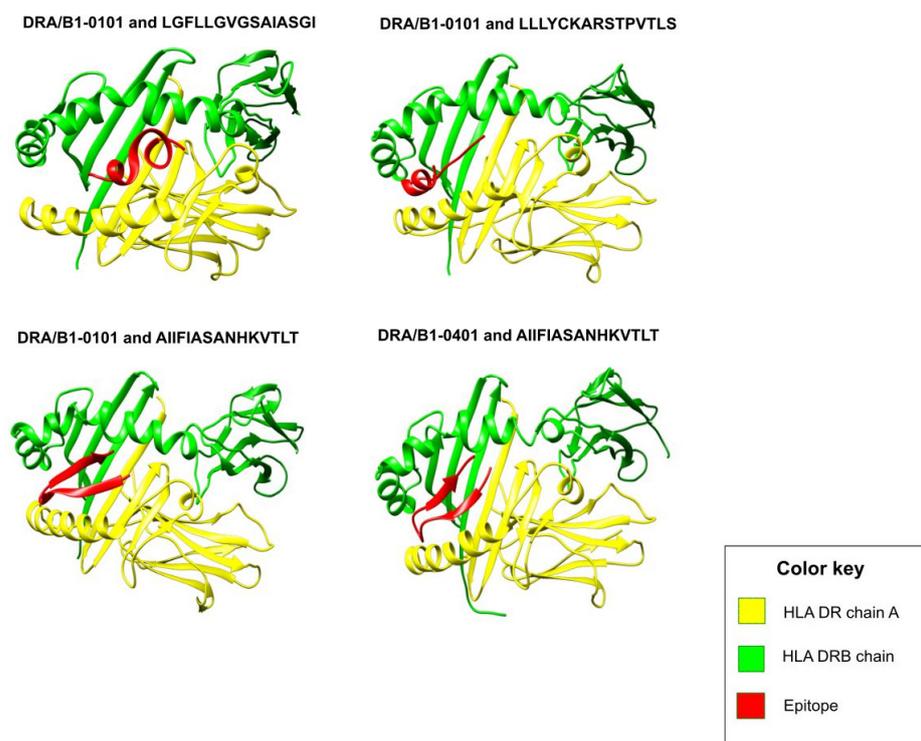
Figure 2. Figure showing structural interactions between the prioritized HLA II epitopes (colored in red) and cognate HLA II chains (colored in yellow and green). Out of all the HLA II epitopes, AIIFIASANHKVTLT (from the G glycoprotein) showed the best HADDOCK scores of −117.4 and −117.1 with HLA DRA/B1*0101 and HLA DRA/B1*0401, respectively. It can be observed that all the epitopes fit comfortably inside the binding groove of the HLA molecule (formed between the two HLA chains shown in yellow and green color, respectively), and thus are projected to interact strongly with their cognate HLA II allele.

In considering all these results, it is expected that the multi-epitope vaccine meets the physiochemical requirements for production. Antigenicity evaluation by AntigenPro as well as VaxiJen servers revealed that the designed vaccine is antigenic. Allergenicity check by both AllergenFP and AllerTOP suggested the non-allergic tendencies of the vaccine, thus it is unlikely that the vaccine will trigger allergic responses in humans. The proposed vaccine's amino acid sequence is provided within Supplementary Material, File S1.

3.3. Structural Modeling of RSV Multi-Epitope Vaccine

Upon careful analysis of structural models initially obtained by all methods, it was found that the RoseTTAFold server yielded better results. RoseTTAFold generated a total of five models. Out of these, one model (MODEL 05) was selected that was considered to have the best stereochemistry quality by the Ramachandran plot and high ERRAT scores (Table 3). However, it was observed that molecular refinements by the Galaxy Refine server did not improve the overall stereochemistry of the vaccine structure. Thus, MODEL 05 was shortlisted for downstream structural analysis. Figure 3 shows the structural evaluation results of MODEL 05. These evaluations were useful to gain crucial insights into structural

modeling output. The best structural model was needed for continued structural analysis, including molecular docking, molecular dynamics simulations and so on. The 3D structure of the multi-epitope vaccine can be visualized in Figure 4.

Table 3. Table showing structural features of models generated by the Robetta server.

Model No.	Ramachandran Plot Analysis			Outlier Residues	ERRAT Score
	Most Favored Region	Additionally Allowed Regions	Generously Allowed Regions		
Model 1	89%	9%	1%	1%	89.9194
Model 2	84.50%	13.50%	1%	1%	96.6942
Model 3	87.50%	11%	0%	1.50%	91.5323
Model 4	89.50%	8%	1%	1.50%	90.1639
Model 5	89%	9.50%	1.50%	0%	91.4286
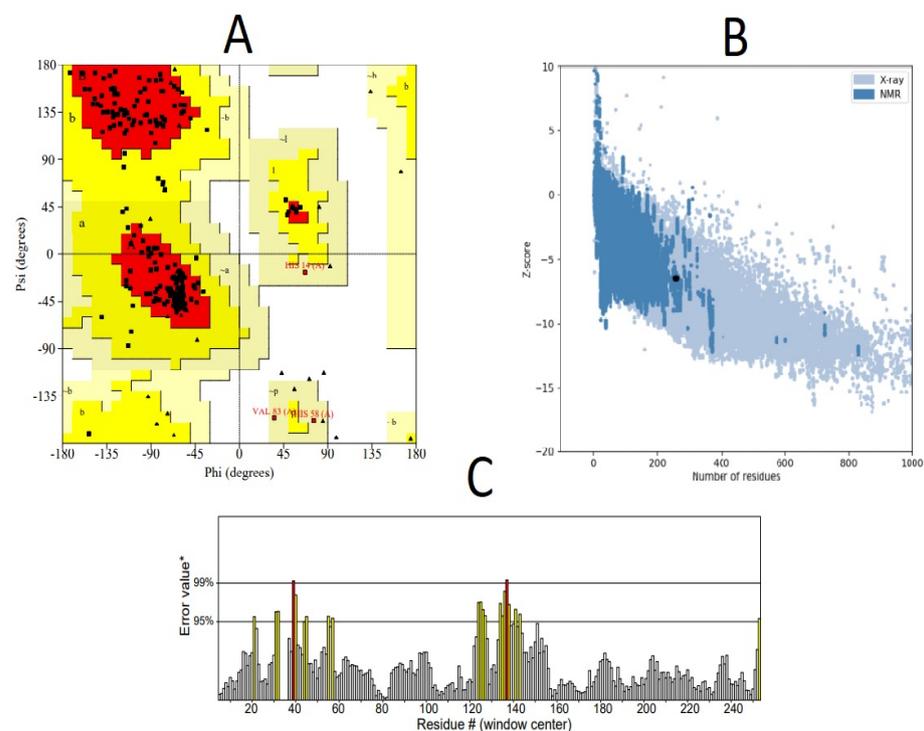
Figure 3. Structural evaluations of the modeled multi-epitope vaccine confirmed its high quality and stability using (**A**) Ramachandran plot analysis. It was confirmed that a total of 89% of the modeled residues resided in the most favored red region (indicated by capital letters A, B) whereas 9.5% and 1.5% residues resided in the additionally allowed bright yellow regions (indicated by small letters a, b) and generously allowed light yellow regions (indicated by ~a, ~b) respectively. Outlier regions are those having white background color. Glycine residues are indicated by triangles. (**B**) Prosa quality check confirmed that the modeled structure's quality lies within the range of similar experimentally resolved NMR structures, with a Z-score of −6.47. (**C**) Stereochemical check by the ERRAT server revealed a high ERRAT score of 91.4286 that signifies that a good quality refined structure was obtained. On the error axis, two lines are drawn to indicate the confidence with which we can reject the regions that exceed that exceed that error value. Regions that have greater than 95% error are indicated by yellow color, whereas in rare cases residues having greater than 99% error values were found, indicated by red color.

(A)

(B)
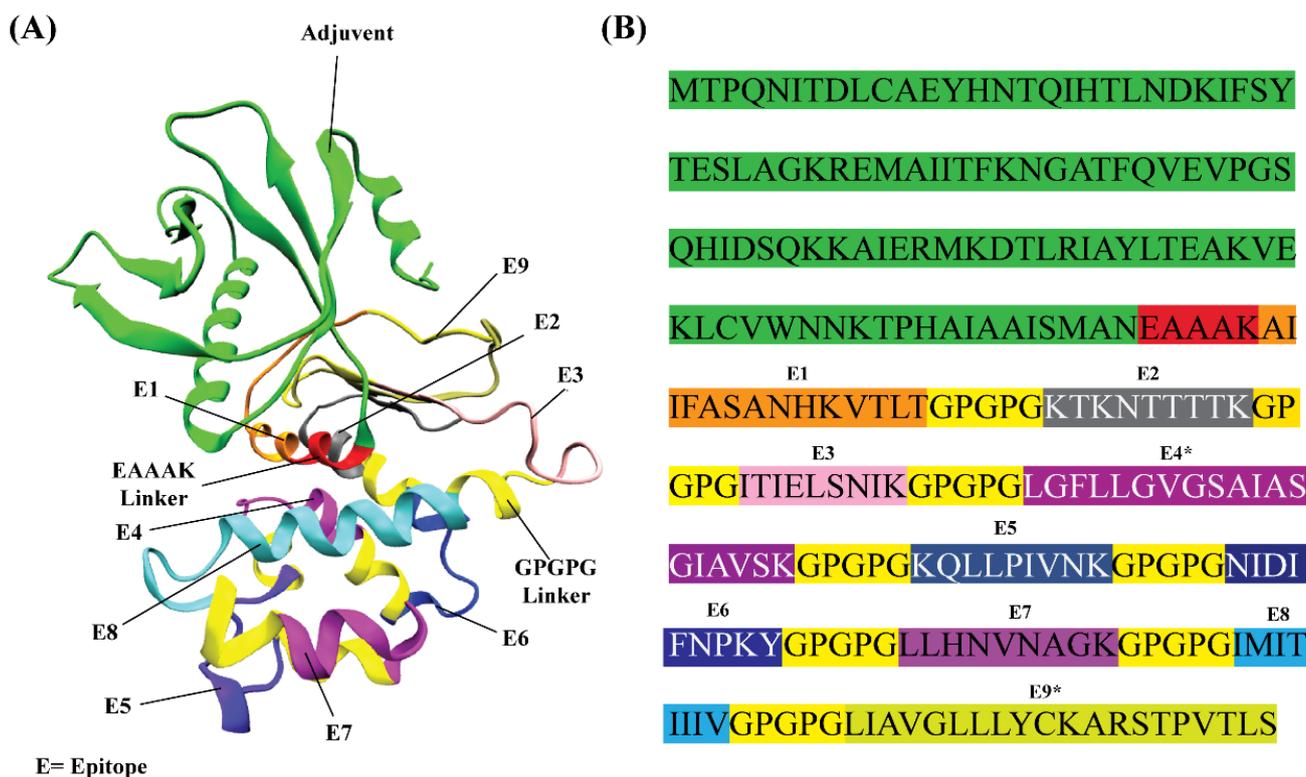
Figure 4. Constructed vaccine for RSV in the present study. (**A**) Modeled three-dimensional structure of vaccine construct. (**B**) Complete amino acid sequence of the vaccine construct. An adjuvant, cholera toxin subunit B, was placed towards the N-terminal of the construct. Specialized linkers (EAAAK and GPGPG) were added to separate the prioritized epitopes. Epitopes 4 and 9 are the merged CD4+ T-cell and CD8+ T-cell epitopes. These two epitopes are indicated by E4* and E9* (stars are added for easy identification).

3.4. Interaction Analysis of Vaccine with Toll-Like Receptor 4 by Molecular Docking

By analyzing docking results of vaccine and TLR4 structures, it was found that the most reputable cluster showed a HADDOCK score of 21.7 ± 4.3. After applying molecular refinements on the representative structure of this complex and clustering, a rapid improvement in the HADDOCK score was noticed. The refined vaccine-TLR4 complex showed a strikingly high HADDOCK score of −143.3 ± 9.6. A negative HADDOCK score signifies a greater potential of favorable interaction between the proteins. The statistical properties of this cluster, along with the HADDOCK score, are provided in Supplementary Material, Table S2. A representative structure of vaccine-TLR4 was visualized, and interacting residues were colored. According to PDBsum's detailed structural characterization, a total of 21 residues of vaccine formed associations with 25 residues of TLR4, with a total of 135 non-bonded contacts, 13 hydrogen bonds and four salt bridges. These interacting chains and interesting molecular interactions can be visualized clearly in Figure 5.
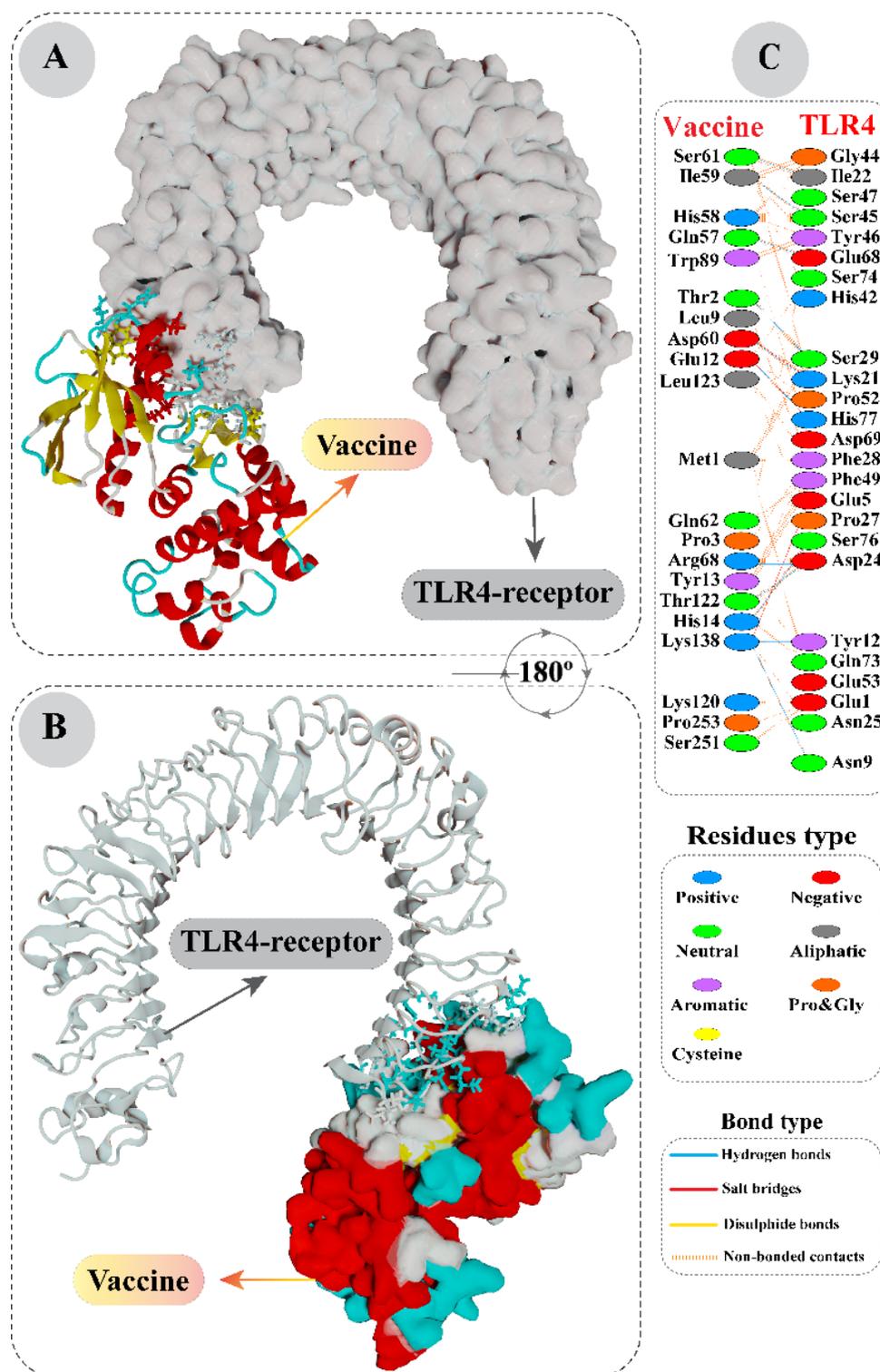
Figure 5. Protein-protein interaction of vaccine and TLR4. (**A**) Refined complex of multi-epitope with TLR4. Here, vaccine structure is shown in ribbon form, whereas the TLR4 structure surface has been represented. (**B**) Same complex as before, but after a 180-degree shift. However, TLR4 structure is shown in ribbon form, whereas the vaccine structure's surface has been represented. (**C**) Structural analysis revealed numerous non-bonded contacts between vaccine and TLR4, as well as some notable hydrogen bonds and salt bridges. The type of amino acid residues involved in interactions have also been noted.

A computational assessment based on Gibbs free energy was conducted on the vaccine-TLR4 complex obtained after docking refinement. This provided crucial information about the binding affinity between the two protein chains. Through this analysis, it was revealed that the docked complex had a -11.3 kcal mol^{-1} ΔG value. The negative sign in Gibb's free energy results strongly indicates that the association between multi-epitope and TLR4 is biologically feasible in terms of thermodynamics. Moreover, the dissociation constant Kd (M) value at 25.0 °C was calculated and found to be 5.6×10^{-9}.

3.5. Molecular Dynamics Analysis on Vaccine-TLR4 Complex

The GROMACS tool analyzed the stability of TLR4 and the multi-epitope ligand complex. A total of 50,000 energy minimization steps were conducted till the system reached <1000 kJ/mol maximum force. Potential energy (E_{pot}) was calculated. Ideally, E_{pot} value should be negative and in the range 10^5–10^6 for protein simulation in water. It was found that the E_{pot} value is -4.42718×10^6 kJ mol^{-1} nm^{-1}, whereas the 4370 atom faced the maximum force of 9.7368×10^2 kJ mol^{-1} nm^{-1} (Figure 6A). Upon heating, the system's temperature rapidly approached 300 K and showed minute fluctuations overall during equilibration (Supplementary Material, Figure S1). NPT equilibration was performed, and it was found that the average value of pressure was 1.6857 (Supplementary Material, Figure S2). The system's average density was calculated to be 1014.57 kg/m^3 as per NPT analysis (Supplementary Material, Figure S3). Mild fluctuations were observed in the protein backbones-based RMSD plot, indicating stability interactions between the two chains (Figure 6B). At 20 ns, the RMSD value approached ~0.4 nm. The RMSF plot (Figure 6C) associated with the protein side-chains showed various peaks, which indicated that many residues within the vaccine-TLR4 complex were highly flexible. The radius of gyration plot (Figure 6D), meanwhile, showed some fluctuations but mainly remained between 3.7 and 4.0 nm throughout the 20-nanosecond simulation. Altogether, these investigations supported a high level of structural stability within the vaccine-TLR complex.
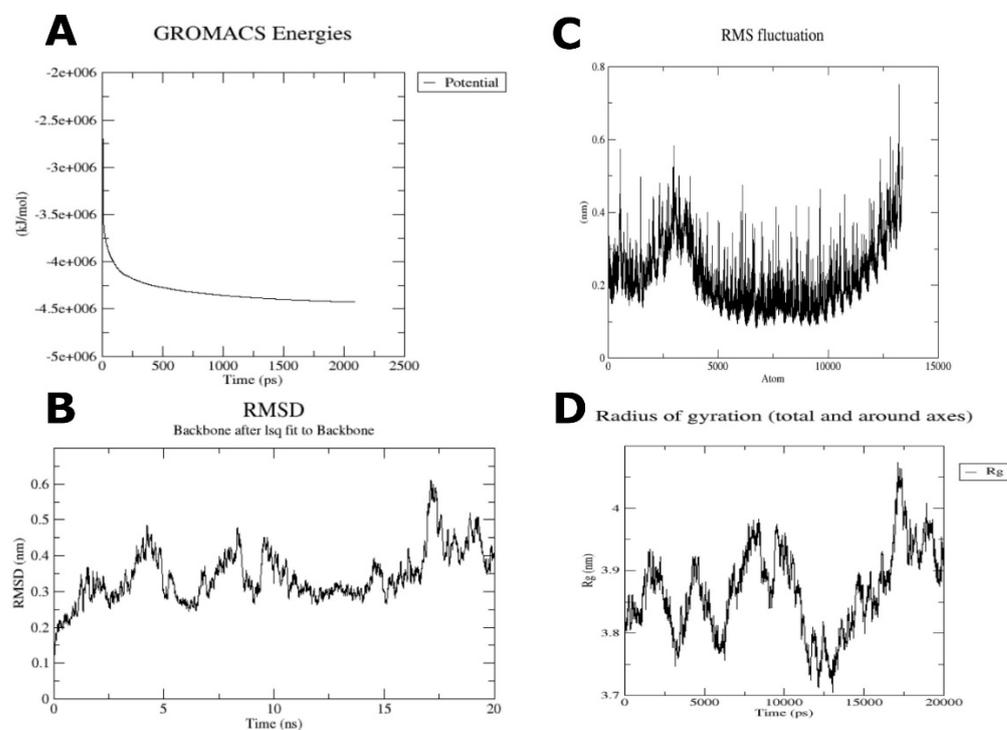
Figure 6. Some graphs obtained after molecular dynamics simulation on the vaccine-TLR4 complex. (**A**) Potential energy during simulation. (**B**) Comparison of backbone RMSD during MD production simulation. (**C**) RMSF of side chains during MD production simulation. (**D**) Radius of gyration with minimum fluctuations, showing compactness of structure during MD production simulation.

3.6. Immune Simulation Analysis

After three in silico doses of the vaccine, various graphs were obtained (Figure 7). Simulated results showed that a high concentration of IgM was noted as a primary response. In addition, the significantly higher level of antibodies activities, including IgM, IgG1 + IgG2, as well as IgM + IgG antibodies in secondary and tertiary responses, were observed. Consequently, rapid clearance of antigen (lower concentration of antigen) was observed (Figure 7A). The vaccine structure induced a robust, long-established immunological response, as seen by the increased quantities of simulated B cells and memory B cell development (Supplementary Material, Figure S4A,B). T-cell numbers, on the other hand, progressively declined following their initial rise with relative immunity development, which is required to elicit immunological responses (Figure 7B,C and Figure S4C). Following the vaccine injections, macrophage activity improved, whereas dendritic cell activity remained stable (Supplementary Material, Figure S4D,E). Similarly, higher levels of cytokine IFN-ϒ (>400,000 ng/mL) were noted (Figure 7D). The elicited immune features suggest that the vaccine could induce highly specific immune responses needed to fight the viral infection.
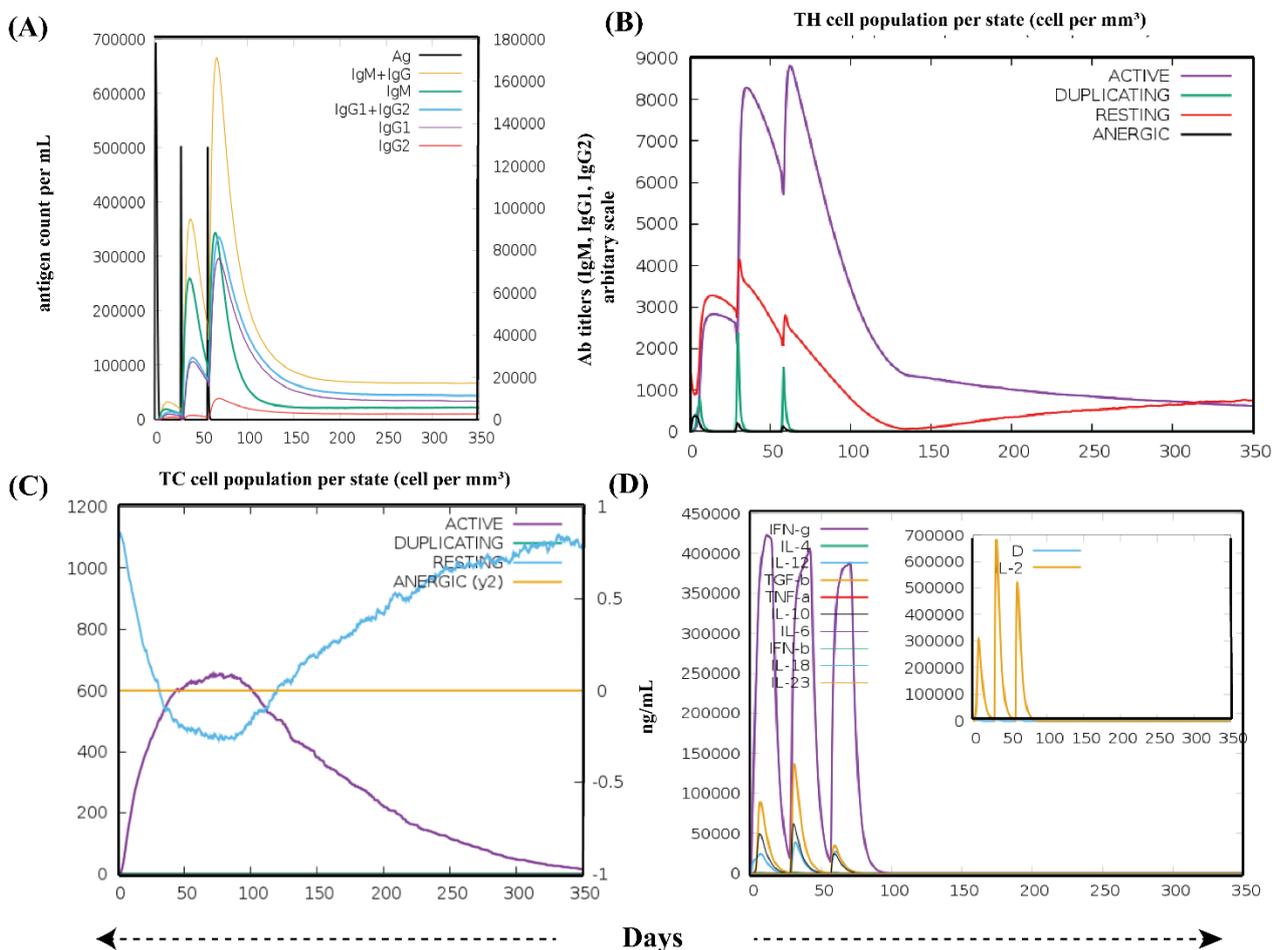

Figure 7. The immunostimulatory potential of the constructed vaccine after three injections of it. (**A**) Immunoglobulin production; (**B**) population/state of helper T-cell; (**C**) population/state of cytotoxic T-cells; (**D**) cytokines and interleukins production (smaller graph) with Simpson index. All units are stated in terms of cells per mm^3 in three consecutive immunological responses.

3.7. Optimized Cloning of the RSV Vaccine Candidate

Upon optimizing the codon sequence of the RSV vaccine, the length of the cDNA sequence was found to be 771 nucleotides (Supplementary Material, File S2). The obtained CAI (0.97) revealed that the adapted sequence is composed of codons suitable for using

the target organism's cellular machinery. Besides, the GC content (52.39%) after codon optimization was found to be acceptable. Therefore, ease of expression is guaranteed inside the *E. coli* K12 strain. Towards the N and C termini of codon-optimized nucleotides, the restriction sites of *Xho*I and *Nde*I enzymes were conjugated, respectively (Figure 8). In order to clone the vaccine, the final optimized sequence was placed within the pET28a (+) vector between the *Xho*I (158) and *Nde*I (936) positions. The cloned plasmid had a final length of 6607 bp.
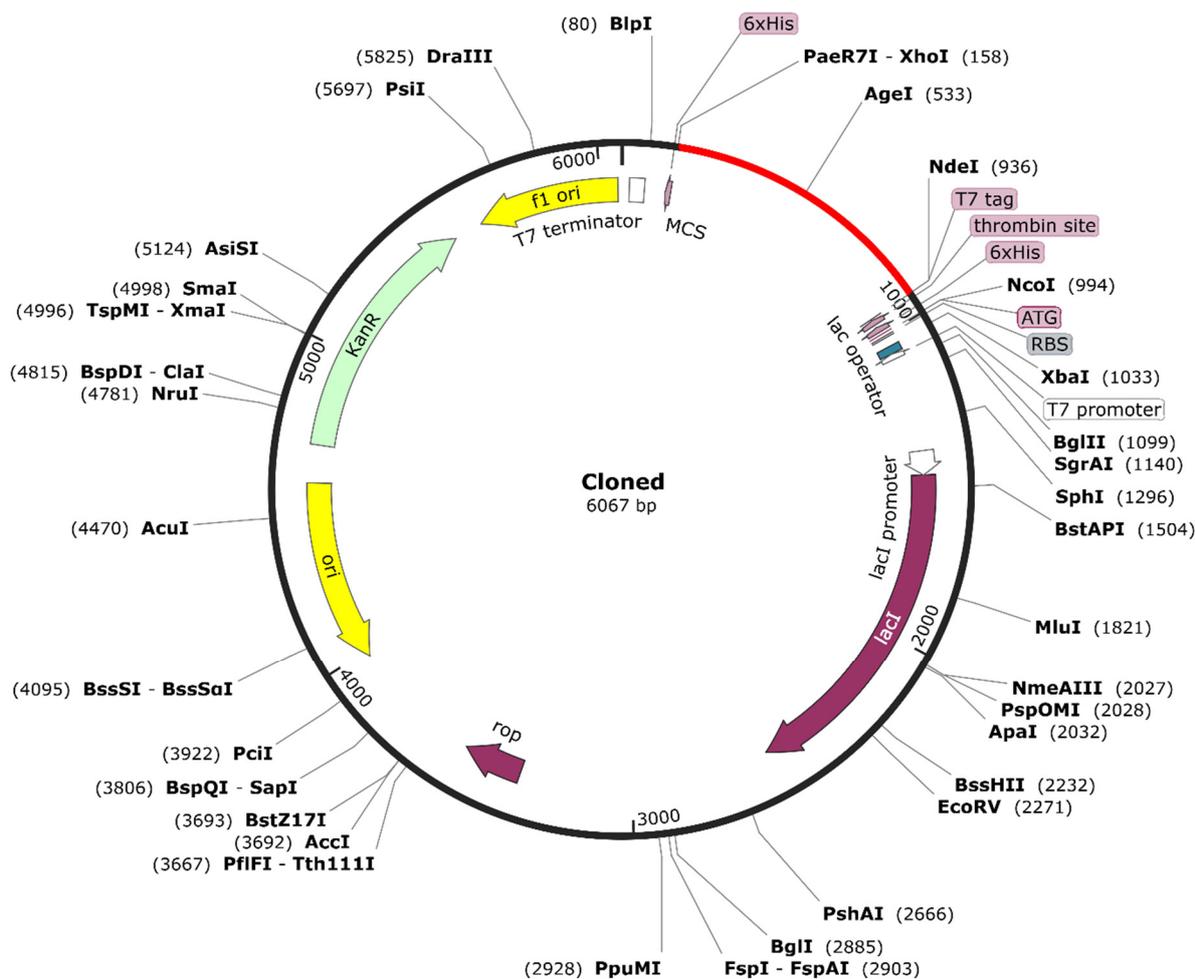
Figure 8. Simulated restriction cloning of the constructed multi-epitope vaccine. The multi-epitope vaccine construct's codon-optimized sequence (shown in red color) was cloned in the pET-28a (+) expression vector (shown in black color) between the *Xho*I (158) and *Nde*I (936) restriction enzyme loci using SnapGene software. For efficient vaccine production, the designed construct can be expressed in *E. coli* (strain K12).

4. Discussion

Vaccination is the perfect health strategy to resolve the menace of infectious diseases, as evidenced by the eradication of the smallpox disease. In this study, our research group explored the concept of an epitope-based anti-RSV vaccine, as opposed to the use of the whole pathogen. Subunit-based vaccines only comprise antigenic regions of the virulent pathogen; thus, they are considered safe for eliciting the appropriate immune responses [77]. Recent progress made in immunoinformatics has enabled the development of vaccines with limited time and resources, which is an encouraging development. However, the success of these computer-guided analyses is dependent on the accurate identification of good antigens and their molecular determinants (epitopes), as well as the efficiency of

the delivery system. Altogether, the identification of good vaccine targets is essential in vaccine studies.

Vaccines that are based on epitopes are known to show good responses. The approach has been followed by other researchers to target various diseases, ranging from cancer to viral and other pathogen-related diseases [21,78–80]. At the forefront of these technologies are numerous immunoinformatics methods and tools that have significantly aided the proper identification of T-cell interacting epitopes. Vaccines based on T-cell epitopes are known to evoke durable immune responses and overcome the antigenic drift that is typically associated with evasion of antibody memory responses [26]. The T-cell subtypes (CD8+ and CD4+) play an important role in triggering antiviral responses and B-cell clonal expansion. Thus, we focused exclusively on the identification of reliable T-cell epitopes, using both sequence-based as well as structural bioinformatics analysis. We also designed a structurally stable multi-epitope vaccine using top-ranked T-cell epitopes, and elaborated on the vaccine's effectiveness using molecular docking and molecular dynamics simulation with TLR4.

Epitope-based vaccines have drawn significant attention over classical immunization because of their several advantages [81]. First, the distinctive design strategy differentiates the epitope-based vaccines from single-epitope or conventional vaccines: they include multiple MHC-restricted epitopes that T-cell receptors can recognize from various T-cell subsets. They can stimulate simultaneous cellular and humoral immune responses due to the inclusion of B-cell and T-cell (CTL and HTL) epitopes. Some special substances, i.e., adjuvants, are added to increase immunogenicity and produce a long-lasting immunological response. Finally, undesirable substances are excluded, thus decreasing the chance of evoking adverse or harmful immune reactions [82]. Therefore, with such benefits, elegant epitope-based vaccines should progress to promising therapeutic options for viral infection. Using the immunoinformatics approach, the multi-epitope vaccine has been designed in recent years for several viruses, including chikungunya, dengue, hepatitis C, MERS-CoV, SARS-CoV-2 and Zika [81,83,84]. Also, several multi-epitope vaccines have already progressed to clinical trials [84]. Molecular dynamics analysis also enabled us to understand the conformational dynamics of the vaccine. It has been reported that regions of increased conformational order are thought to be more antigenic [85]. Considering the significant health burden of RSV infection, a prophylactic vaccine against the virus is desirable.

In this study, we aimed to assess the antigenic potential of two very important glycoproteins of RSV: fusion F and attachment G. Due to their role in viral attachment and entry, they are deemed to be crucial for RSV pathogenesis [10]. On top of that, they are the only RSV proteins known to target the human cell membrane and contain conserved regions. Therefore, we considered it prudent to choose them as targets for developing a potent anti-RSV vaccine.

Adverse effects may occur after administering vaccines, and could be due to any vaccine constituents [86,87]. The probability of generating harmful autoimmune responses cannot be ruled out [88]. For example, administering of hepatitis B vaccine may lead to thrombocytopenia and/or polyarthritis [89]. Similarly, vaccines against tetanus and rabies diseases could potentially cause neural complications [90]. In this study, the prioritized epitopes were checked through BLASTp analysis against the human genome, in order to determine whether they have autoimmune potential or not. Our analyses suggested that these epitopes are not anticipated to trigger autoimmunity, thus rendering them safe for use.

While bioinformatics tools are known to yield good results, one major issue faced by researchers is that the number of epitopes obtained after in silico filters is often quite high [91]. It is advisable to prioritize only the best and most promising epitopes out of the pool of epitopes obtained so that other scientists can validate these results by wet laboratory approaches. In order to reduce the epitope list to a manageable level, we applied filters such as antigenicity, non-homology to humans, immunogenicity and good HLA binding affinity. Moreover, we explored the structural associations of these finalized T-cell epitopes

with their cognate HLA allele(s) using molecular docking. For this purpose, we used the HADDOCK 2.2 server, a highly robust and effective tool to conduct protein-peptide docking. Docking results revealed a strong association between the epitopes and their corresponding HLA allele(s), which further validated our approach.

Prioritization of epitopes by adopting an in silico sequence-based characterization and molecular docking analysis is a significant step towards the development of peptide-based vaccine candidates. However, more follow-up confirmatory experiments are crucial to validate the effectiveness of the epitopes as well as the multi-epitope vaccine against RSV. Nevertheless, the immunological data obtained in this study can be integrated with data from other databases related to genomics, proteomics or pharmacogenomics. By doing this, the scope and breadth of immunoinformatics analysis presented in this study can be increased.

5. Conclusions

In this study, our research group adopted a specialized immunoinformatics workflow to develop a multi-epitope subunit vaccine containing highly antigenic T-cell epitopes of the RSV glycoproteins F and G. Elaborate investigations using molecular docking, thermodynamic investigations, molecular dynamics and in silico simulation analyses strongly support the ability of our vaccine to not only associate strongly with Toll-like receptor 4, but also to trigger highly specific cellular and humoral immune responses against RSV pathogens. Altogether, the present study shows that bioinformatics approaches can help develop effective treatments for emerging viral infections, despite limited resources and time. However, these results are purely computational and need to be corroborated by wet laboratory methods.

References

1. Elhalik, M.; El-Atawi, K.; Dash, S.K.; Faquih, A.; Satyan, A.D.; Gourshettiwar, N.; Khan, A.; Varughese, S.; Ramesh, A.; Khamis, E. Palivizumab Prophylaxis among Infants at Increased Risk of Hospitalization Due to Respiratory Syncytial Virus Infection in UAE: A Hospital-Based Study. *Can. Respir. J.* **2019**, *2019*, 2986286. [CrossRef]

2. Griffiths, C.; Drews, S.J.; Marchant, D.J. Respiratory Syncytial Virus: Infection, Detection, and New Options for Prevention and Treatment. *Clin. Microbiol. Rev.* **2017**, *30*, 277–319. [CrossRef]

3. Afonso, C.L.; Amarasinghe, G.K.; Bányai, K.; Bào, Y.; Basler, C.F.; Bavari, S.; Bejerman, N.; Blasdell, K.R.; Briand, F.-X.; Briese, T.; et al. Taxonomy of the Order Mononegavirales: Update 2016. *Arch. Virol.* **2016**, *161*, 2351–2360. [CrossRef]

4. ul Qamar, M.; Shokat, Z.; Muneer, I.; Ashfaq, U.A.; Javed, H.; Anwar, F.; Bari, A.; Zahid, B.; Saari, N. Multiepitope-Based Subunit Vaccine Design and Evaluation against Respiratory Syncytial Virus Using Reverse Vaccinology Approach. *Vaccines* **2020**, *8*, 288. [CrossRef]

5. Mufson, M.A.; Örvell, C.; Rafnar, B.; Norrby, E. Two Distinct Subtypes of Human Respiratory Syncytial Virus. *J. Gen. Virol.* **1985**, *66*, 2111–2124. [CrossRef]

6. Thongpan, I.; Mauleekoonphairoj, J.; Vichiwattana, P.; Korkong, S.; Wasitthankasem, R.; Vongpunsawad, S.; Poovorawan, Y. Respiratory Syncytial Virus Genotypes NA1, ON1, and BA9 Are Prevalent in Thailand, 2012–2015. *PeerJ* **2017**, *5*, e3970. [CrossRef]

7. Collins, P.L.; Fearns, R.; Graham, B.S. Respiratory Syncytial Virus: Virology, Reverse Genetics, and Pathogenesis of Disease. In *Challenges and Opportunities for Respiratory Syncytial Virus Vaccines*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 3–38.

8. Carvajal, J.J.; Avellaneda, A.M.; Salazar-Ardiles, C.N.; Maya, J.E.; Kalergis, A.M.; Lay, M.K.-L. Host Components Contributing to Respiratory Syncytial Virus Pathogenesis. *Front. Immunol.* **2019**, *10*, 2152. [CrossRef]

9. Mastrangelo, P.; Hegele, R.G. RSV Fusion: Time for a New Model. *Viruses* **2013**, *5*, 873–885. [CrossRef]

10. McLellan, J.S.; Ray, W.C.; Peeples, M.E. Structure and Function of Respiratory Syncytial Virus Surface Glycoproteins. In *Challenges and Opportunities for Respiratory Syncytial Virus Vaccines*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 83–104.

11. Naz, R.; Gul, A.; Javed, U.; Urooj, A.; Amin, S.; Fatima, Z. Etiology of Acute Viral Respiratory Infections Common in Pakistan: A Review. *Rev. Med. Virol.* **2019**, *29*, e2024. [CrossRef]

12. Goins, W.P.; Talbot, H.K.; Talbot, T.R. Health Care—Acquired Viral Respiratory Diseases. *Infect. Dis. Clin.* **2011**, *25*, 227–244. [CrossRef]

13. Schmidt, M.R.; McGinnes, L.W.; Kenward, S.A.; Willems, K.N.; Woodland, R.T.; Morrison, T.G. Long-Term and Memory Immune Responses in Mice against Newcastle Disease Virus-like Particles Containing Respiratory Syncytial Virus Glycoprotein Ectodomains. *J. Virol.* **2012**, *86*, 11654–11662. [CrossRef]

14. Schweitzer, J.W.; Justice, N.A. Respiratory Syncytial Virus Infection. 2017. Available online: https://www.ncbi.nlm.nih.gov/books/NBK459215/ (accessed on 3 July 2022).

15. Falsey, A.R.; McCann, R.M.; Hall, W.J.; Criddle, M.M.; Formica, M.A.; Wycoff, D.; Kolassa, J.E. The "Common Cold" in Frail Older Persons: Impact of Rhinovirus and Coronavirus in a Senior Daycare Center. *J. Am. Geriatr. Soc.* **1997**, *45*, 706–711. [CrossRef]

16. Saravolatz, L.D.; Empey, K.M.; Peebles Jr, R.S.; Kolls, J.K. Pharmacologic Advances in the Treatment and Prevention of Respiratory Syncytial Virus. *Clin. Infect. Dis.* **2010**, *50*, 1258–1267.

17. IMpact-RSV Study Group. Palivizumab, a Humanized Respiratory Syncytial Virus Monoclonal Antibody, Reduces Hospitalization from Respiratory Syncytial Virus Infection in High-Risk Infants. *Pediatrics* **1998**, *102*, 531–537. [CrossRef]

18. Ventre, K.; Randolph, A. Ribavirin for Respiratory Syncytial Virus Infection of the Lower Respiratory Tract in Infants and Young Children. *Cochrane Database Syst. Rev.* **2007**, *1*, CD000181.

19. Hall, C.B.; McBride, J.T.; Walsh, E.E.; Bell, D.M.; Gala, C.L.; Hildreth, S.; Ten Eyck, L.G.; Hall, W.J. Aerosolized Ribavirin Treatment of Infants with Respiratory Syncytial Viral Infection: A Randomized Double-Blind Study. *N. Engl. J. Med.* **1983**, *308*, 1443–1447. [CrossRef]

20. Sette, A.; Rappuoli, R. Reverse Vaccinology: Developing Vaccines in the Era of Genomics. *Immunity* **2010**, *33*, 530–541. [CrossRef]

21. Dar, H.A.; Zaheer, T.; Shehroz, M.; Ullah, N.; Naz, K.; Muhammad, S.A.; Zhang, T.; Ali, A. Immunoinformatics-Aided Design and Evaluation of a Potential Multi-Epitope Vaccine against Klebsiella Pneumoniae. *Vaccines* **2019**, *7*, 88. [CrossRef]

22. Alberts, B.; Johnson, A.; Lewis, J.; Raff, M.; Roberts, K.; Walter, P. T Cells and MHC Proteins. In *Molecular Biology of the Cell*, 4th ed.; Garland Science: New York, NY, USA, 2002.

23. Ayub, G.; Waheed, Y.; Najmi, M.H. Prediction and Conservancy Analysis of Promiscuous T-Cell Binding Epitopes of Ebola Virus L Protein: An in Silico Approach. *Asian Pacific J. Trop. Dis.* **2016**, *6*, 169–173. [CrossRef]

24. Oyarzun, P.; Ellis, J.J.; Gonzalez-Galarza, F.F.; Jones, A.R.; Middleton, D.; Boden, M.; Kobe, B. A Bioinformatics Tool for Epitope-Based Vaccine Design That Accounts for Human Ethnic Diversity: Application to Emerging Infectious Diseases. *Vaccine* **2015**, *33*, 1267–1273. [CrossRef]

25. Sette, A.; Fikes, J. Epitope-Based Vaccines: An Update on Epitope Identification, Vaccine Design and Delivery. *Curr. Opin. Immunol.* **2003**, *15*, 461–470. [CrossRef]

26. Patronov, A.; Doytchinova, I. T-Cell Epitope Vaccine Design by Immunoinformatics. *Open Biol.* **2013**, *3*, 120139. [CrossRef]

27. Barh, D.; Barve, N.; Gupta, K.; Chandra, S.; Jain, N.; Tiwari, S.; Leon-Sicairos, N.; Canizalez-Roman, A.; dos Santos, A.R.; Hassan, S.S.; et al. Exoproteome and Secretome Derived Broad Spectrum Novel Drug and Vaccine Candidates in Vibrio Cholerae Targeted by Piper Betel Derived Compounds. *PLoS ONE* **2013**, *8*, e52773. [CrossRef]

28. Compton, T.; Kurt-Jones, E.A.; Boehme, K.W.; Belko, J.; Latz, E.; Golenbock, D.T.; Finberg, R.W. Human Cytomegalovirus Activates Inflammatory Cytokine Responses via CD14 and Toll-like Receptor 2. *J. Virol.* **2003**, *77*, 4588–4596. [CrossRef]

29. Vijay, K. Toll-like Receptors in Immunity and Inflammatory Diseases: Past, Present, and Future. *Int. Immunopharmacol.* **2018**, *59*, 391–412. [CrossRef]

30. Kerepesi, L.A.; Leon, O.; Lustigman, S.; Abraham, D. Protective Immunity to the Larval Stages of Onchocerca Volvulus Is Dependent on Toll-like Receptor 4. *Infect. Immun.* **2005**, *73*, 8291–8297. [CrossRef]

31. Andreatta, M.; Nielsen, M. Gapped Sequence Alignment Using Artificial Neural Networks: Application to the MHC Class I System. *Bioinformatics* **2015**, *32*, 511–517. [CrossRef]

32. Nielsen, M.; Lundegaard, C.; Worning, P.; Lauemøller, S.L.; Lamberth, K.; Buus, S.; Brunak, S.; Lund, O. Reliable Prediction of T-Cell Epitopes Using Neural Networks with Novel Sequence Representations. *Protein Sci.* **2003**, *12*, 1007–1017. [CrossRef]

33. Doytchinova, I.A.; Flower, D.R. VaxiJen: A Server for Prediction of Protective Antigens, Tumour Antigens and Subunit Vaccines. *BMC Bioinform.* **2007**, *8*, 4. [CrossRef]

34. Calis, J.J.A.; Maybeno, M.; Greenbaum, J.A.; Weiskopf, D.; De Silva, A.D.; Sette, A.; Keşmir, C.; Peters, B. Properties of MHC Class I Presented Peptides That Enhance Immunogenicity. *PLoS Comput. Biol.* **2013**, *9*, e1003266. [CrossRef]

35. Guan, P.; Doytchinova, I.A.; Zygouri, C.; Flower, D.R. MHCPred: A Server for Quantitative Prediction of Peptide–MHC Binding. *Nucleic Acids Res.* **2003**, *31*, 3621–3624. [CrossRef]

36. Jensen, K.K.; Andreatta, M.; Marcatili, P.; Buus, S.; Greenbaum, J.A.; Yan, Z.; Sette, A.; Peters, B.; Nielsen, M. Improved Methods for Predicting Peptide Binding Affinity to MHC Class II Molecules. *Immunology* **2018**, *154*, 394–406. [CrossRef]

37. Chen, H.; Zhou, H.-X. Prediction of Solvent Accessibility and Sites of Deleterious Mutations from Protein Sequence. *Nucleic Acids Res.* **2005**, *33*, 3193–3199. [CrossRef]

38. 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–W454. [CrossRef]

39. van Zundert, G.C.P.; Rodrigues, J.P.G.L.M.; Trellet, M.; Schmitz, C.; Kastritis, P.L.; Karaca, E.; Melquiond, A.S.J.; van Dijk, M.; de Vries, S.J.; Bonvin, A.M.J.J. The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes. *J. Mol. Biol.* **2016**, *428*, 720–725. [CrossRef]

40. De Vries, S.J.; Van Dijk, M.; Bonvin, A.M.J.J. The HADDOCK Web Server for Data-Driven Biomolecular Docking. *Nat. Protoc.* **2010**, *5*, 883. [CrossRef]

41. Gerstner, C.; Dubnovitsky, A.; Sandin, C.; Kozhukh, G.; Uchtenhagen, H.; James, E.A.; Rönnelid, J.; Ytterberg, A.J.; Pieper, J.; Reed, E.; et al. Functional and Structural Characterization of a Novel HLA-DRB1* 04: 01-Restricted α-Enolase T Cell Epitope in Rheumatoid Arthritis. *Front. Immunol.* **2016**, *7*, 494. [CrossRef]

42. Quiñones-Parra, S.; Grant, E.; Loh, L.; Nguyen, T.H.O.; Campbell, K.-A.; Tong, S.Y.C.; Miller, A.; Doherty, P.C.; Vijaykrishna, D.; Rossjohn, J.; et al. Preexisting CD8+ T-Cell Immunity to the H7N9 Influenza A Virus Varies across Ethnicities. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 1049–1054. [CrossRef]

43. Zhang, S.; Liu, J.; Cheng, H.; Tan, S.; Qi, J.; Yan, J.; Gao, G.F. Structural Basis of Cross-Allele Presentation by HLA-A* 0301 and HLA-A* 1101 Revealed by Two HIV-Derived Peptide Complexes. *Mol. Immunol.* **2011**, *49*, 395–401. [CrossRef]

44. Li, L.; Bouvier, M. Structures of HLA-A* 1101 Complexed with Immunodominant Nonamer and Decamer HIV-1 Epitopes Clearly Reveal the Presence of a Middle, Secondary Anchor Residue. *J. Immunol.* **2004**, *172*, 6175–6184. [CrossRef]

45. Murthy, V.L.; Stern, L.J. The Class II MHC Protein HLA-DR1 in Complex with an Endogenous Peptide: Implications for the Structural Basis of the Specificity of Peptide Binding. *Structure* **1997**, *5*, 1385–1396. [CrossRef]

46. Khan, A.R.; Baker, B.M.; Ghosh, P.; Biddison, W.E.; Wiley, D.C. The Structure and Stability of an HLA-A* 0201/Octameric Tax Peptide Complex with an Empty Conserved Peptide-N-Terminal Binding Site. *J. Immunol.* **2000**, *164*, 6398–6405. [CrossRef]

47. Laskowski, R.A. PDBsum: Summaries and Analyses of PDB Structures. *Nucleic Acids Res.* **2001**, *29*, 221–222. [CrossRef]

48. Pettersen, E.F.; Goddard, T.D.; Huang, C.C.; Couch, G.S.; Greenblatt, D.M.; Meng, E.C.; Ferrin, T.E. UCSF Chimera—A Visualization System for Exploratory Research and Analysis. *J. Comput. Chem.* **2004**, *25*, 1605–1612. [CrossRef]

49. Gasteiger, E.; Gattiker, A.; Hoogland, C.; Ivanyi, I.; Appel, R.D.; Bairoch, A. ExPASy: The Proteomics Server for in-Depth Protein Knowledge and Analysis. *Nucleic Acids Res.* **2003**, *31*, 3784–3788. [CrossRef]

50. Magnan, C.N.; Zeller, M.; Kayala, M.A.; Vigil, A.; Randall, A.; Felgner, P.L.; Baldi, P. High-Throughput Prediction of Protein Antigenicity Using Protein Microarray Data. *Bioinformatics* **2010**, *26*, 2936–2943. [CrossRef]

51. Dimitrov, I.; Bangov, I.; Flower, D.R.; Doytchinova, I. AllerTOP v. 2—a Server for in Silico Prediction of Allergens. *J. Mol. Model.* **2014**, *20*, 2278. [CrossRef]

52. Dimitrov, I.; Naneva, L.; Doytchinova, I.; Bangov, I. AllergenFP: Allergenicity Prediction by Descriptor Fingerprints. *Bioinformatics* **2014**, *30*, 846–851. [CrossRef]

53. Kelley, L.A.; Mezulis, S.; Yates, C.M.; Wass, M.N.; Sternberg, M.J.E. The Phyre2 Web Portal for Protein Modeling, Prediction and Analysis. *Nat. Protoc.* **2015**, *10*, 845–858. [CrossRef]

54. Zhang, Y. I-TASSER Server for Protein 3D Structure Prediction. *BMC Bioinform.* **2008**, *9*, 40. [CrossRef]

55. 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*, 1511–1522. [CrossRef] [PubMed]

56. Baek, M.; DiMaio, F.; Anishchenko, I.; Dauparas, J.; Ovchinnikov, S.; Lee, G.R.; Wang, J.; Cong, Q.; Kinch, L.N.; Schaeffer, R.D.; et al. Accurate Prediction of Protein Structures and Interactions Using a Three-Track Neural Network. *Science* **2021**, *373*, 871–876. [CrossRef] [PubMed]

57. Cheng, J.; Randall, A.Z.; Sweredoski, M.J.; Baldi, P. SCRATCH: A Protein Structure and Structural Feature Prediction Server. *Nucleic Acids Res.* **2005**, *33*, W72–W76. [CrossRef]

58. Laskowski, R.A.; MacArthur, M.W.; Moss, D.S.; Thornton, J.M. PROCHECK: A Program to Check the Stereochemical Quality of Protein Structures. *J. Appl. Crystallogr.* **1993**, *26*, 283–291. [CrossRef]

59. Colovos, C.; Yeates, T. ERRAT: An Empirical Atom-Based Method for Validating Protein Structures. *Protein Sci.* **1993**, *2*, 1511–1519. [CrossRef]

60. Wiederstein, M.; Sippl, M.J. ProSA-Web: Interactive Web Service for the Recognition of Errors in Three-Dimensional Structures of Proteins. *Nucleic Acids Res.* **2007**, *35*, W407–W410. [CrossRef]

61. Sippl, M.J. Recognition of Errors in Three-Dimensional Structures of Proteins. *Proteins Struct. Funct. Bioinform.* **1993**, *17*, 355–362. [CrossRef]

62. Heo, L.; Park, H.; Seok, C. GalaxyRefine: Protein Structure Refinement Driven by Side-Chain Repacking. *Nucleic Acids Res.* **2013**, *41*, W384–W388. [CrossRef]

63. Park, B.S.; Song, D.H.; Kim, H.M.; Choi, B.-S.; Lee, H.; Lee, J.-O. The Structural Basis of Lipopolysaccharide Recognition by the TLR4–MD-2 Complex. *Nature* **2009**, *458*, 1191–1195. [CrossRef]

64. de Vries, S.J.; Bonvin, A.M.J.J. CPORT: A Consensus Interface Predictor and Its Performance in Prediction-Driven Docking with HADDOCK. *PLoS ONE* **2011**, *6*, e17695. [CrossRef]

65. Xue, L.C.; Rodrigues, J.P.; Kastritis, P.L.; Bonvin, A.M.; Vangone, A. PRODIGY: A Web Server for Predicting the Binding Affinity of Protein–Protein Complexes. *Bioinformatics* **2016**, *32*, 3676–3678. [CrossRef] [PubMed]

66. Abraham, M.J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J.C.; Hess, B.; Lindahl, E. GROMACS: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **2015**, *1*, 19–25. [CrossRef]

67. Kaminski, G.A.; Friesner, R.A.; Tirado-Rives, J.; Jorgensen, W.L. Evaluation and Reparametrization of the OPLS-AA Force Field for Proteins via Comparison with Accurate Quantum Chemical Calculations on Peptides. *J. Phys. Chem. B* **2001**, *105*, 6474–6487. [CrossRef]

68. Turner, P. *XMGRACE, Version 5.1. 19*; Center for Coastal and Land-Margin Research, Oregon Graduate Institute of Science and Technology: Beaverton, OR, USA, 2005.

69. Rapin, N.; Lund, O.; Bernaschi, M.; Castiglione, F. Computational Immunology Meets Bioinformatics: The Use of Prediction Tools for Molecular Binding in the Simulation of the Immune System. *PLoS ONE* **2010**, *5*, e9862. [CrossRef] [PubMed]

70. Ismail, S.; Ahmad, S.; Azam, S.S. Immunoinformatics Characterization of SARS-CoV-2 Spike Glycoprotein for Prioritization of Epitope Based Multivalent Peptide Vaccine. *J. Mol. Liq.* **2020**, *314*, 113612. [CrossRef] [PubMed]

71. Chauhan, V.; Singh, M.P. Immuno-Informatics Approach to Design a Multi-Epitope Vaccine to Combat Cytomegalovirus Infection. *Eur. J. Pharm. Sci.* **2020**, *147*, 105279. [CrossRef]

72. Madeira, F.; Park, Y.M.; Lee, J.; Buso, N.; Gur, T.; Madhusoodanan, N.; Basutkar, P.; Tivey, A.R.N.; Potter, S.C.; Finn, R.D.; et al. The EMBL-EBI Search and Sequence Analysis Tools APIs in 2019. *Nucleic Acids Res.* **2019**, *47*, W636–W641. [CrossRef]

73. Grote, A.; Hiller, K.; Scheer, M.; Münch, R.; Nörtemann, B.; Hempel, D.C.; 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–W531. [CrossRef]

74. Krishnarjuna, B.; Andrew, D.; MacRaild, C.A.; Morales, R.A.V.; Beeson, J.G.; Anders, R.F.; Richards, J.S.; Norton, R.S. Strain-Transcending Immune Response Generated by Chimeras of the Malaria Vaccine Candidate Merozoite Surface Protein 2. *Sci. Rep.* **2016**, *6*, 20613. [CrossRef]

75. Trimaille, T.; Verrier, B. Micelle-Based Adjuvants for Subunit Vaccine Delivery. *Vaccines* **2015**, *3*, 803–813. [CrossRef]

76. Zaheer, T.; Waseem, M.; Waqar, W.; Dar, H.A.; Shehroz, M.; Naz, K.; Ishaq, Z.; Ahmad, T.; Ullah, N.; Bakhtiar, S.M.; et al. Anti-COVID-19 Multi-Epitope Vaccine Designs Employing Global Viral Genome Sequences. *PeerJ* **2020**, *8*, e9541. [CrossRef]

77. Khatoon, N.; Pandey, R.K.; Prajapati, V.K. Exploring Leishmania Secretory Proteins to Design B and T Cell Multi-Epitope Subunit Vaccine Using Immunoinformatics Approach. *Sci. Rep.* **2017**, *7*, 8285. [CrossRef] [PubMed]

78. Jabbar, B.; Rafique, S.; Salo-Ahen, O.M.H.; Ali, A.; Munir, M.; Idrees, M.; Mirza, M.U.; Vanmeert, M.; Shah, S.Z.; Jabbar, I.; et al. Antigenic Peptide Prediction from E6 and E7 Oncoproteins of HPV Types 16 and 18 for Therapeutic Vaccine Design Using Immunoinformatics and MD Simulation Analysis. *Front. Immunol.* **2018**, *9*, 3000. [CrossRef] [PubMed]

79. Almanzar, G.; Herndler-Brandstetter, D.; Chaparro, S.V.; Jenewein, B.; Keller, M.; Grubeck-Loebenstein, B. Immunodominant Peptides from Conserved Influenza Proteins–A Tool for More Efficient Vaccination in the Elderly? *Wien. Med. Wochenschr.* **2007**, *157*, 116–121. [CrossRef] [PubMed]

80. Oyarzún, P.; Kobe, B. Recombinant and Epitope-Based Vaccines on the Road to the Market and Implications for Vaccine Design and Production. *Hum. Vaccines Immunother.* **2016**, *12*, 763–767. [CrossRef]

81. Dong, R.; Chu, Z.; Yu, F.; Zha, Y. Contriving Multi-Epitope Subunit of Vaccine for COVID-19: Immunoinformatics Approaches. *Front. Immunol.* **2020**, *11*, 1784. [CrossRef]

82. Zhang, L. Multi-Epitope Vaccines: A Promising Strategy against Tumors and Viral Infections. *Cell. Mol. Immunol.* **2018**, *15*, 182–184. [CrossRef]

83. Alam, A.; Khan, A.; Imam, N.; Siddiqui, M.F.; Waseem, M.; Malik, M.Z.; Ishrat, R. Design of an Epitope-Based Peptide Vaccine against the SARS-CoV-2: A Vaccine-Informatics Approach. *Brief. Bioinform.* **2021**, *22*, 1309–1323. [CrossRef]

84. Mahmud, S.; Rafi, M.; Paul, G.K.; Promi, M.M.; Shimu, M.; Sultana, S.; Biswas, S.; Emran, T.B.; Dhama, K.; Alyami, S.A.; et al. Designing a Multi-Epitope Vaccine Candidate to Combat MERS-CoV by Employing an Immunoinformatics Approach. *Sci. Rep.* **2021**, *11*, 15431. [CrossRef]

85. MacRaild, C.A.; Zachrdla, M.; Andrew, D.; Krishnarjuna, B.; Nováček, J.; Žídek, L.; Sklenář, V.; Richards, J.S.; Beeson, J.G.; Anders, R.F.; et al. Conformational Dynamics and Antigenicity in the Disordered Malaria Antigen Merozoite Surface Protein 2. *PLoS ONE* **2015**, *10*, e0119899. [CrossRef]

86. Jain, S.; Baranwal, M. Conserved Peptide Vaccine Candidates Containing Multiple Ebola Nucleoprotein Epitopes Display Interactions with Diverse HLA Molecules. *Med. Microbiol. Immunol.* **2019**, *208*, 227–238. [CrossRef] [PubMed]

87. Fritsche, P.J. Vaccine Hypersensitivity-Update and Overview. *Swiss Med. Wkly.* **2010**, *140*, 238–246. [PubMed]

88. Cohen, A.D.; Shoenfeld, Y. Vaccine-Induced Autoimmunity. *J. Autoimmun.* **1996**, *9*, 699. [CrossRef] [PubMed]

89. Meyboom, R.H.B.; Fucik, H.; Edwards, I.R. Thrombocytopenia Reported in Association with Hepatitis B and A Vaccines. *Lancet* **1995**, *345*, 1638. [CrossRef]

90. Topaloglu, H.; Berker, M.; Kansu, T.; Saatci, U.; Renda, Y. Optic Neuritis and Myelitis after Booster Tetanus Toxoid Vaccination. *Lancet (Br. Ed.)* **1992**, *339*, 178–179. [CrossRef]

91. Akya, A.; Farasat, A.; Ghadiri, K.; Rostamian, M. Identification of HLA-I Restricted Epitopes in Six Vaccine Candidates of Leishmania Tropica Using Immunoinformatics and Molecular Dynamics Simulation Approaches. *Infect. Genet. Evol.* **2019**, *75*, 103953. [CrossRef]