Development of Conserved Multi-epitopes Based Hybrid Vaccine against SARS-CoV-2 Variants; An Immunoinformatic Approach

Allah Rakha Yaseen (✉ 123allah.rakha@gmail.com)

University of the Punjab

Muhammad Suleman

University of the Punjab

Abstract

The world had faced unprecedented disruptions like global quarantine and the COVID-19 pandemic due to SARS-CoV-2. To combat the unsettling situations, several effective vaccines have been developed and are currently being used. However, the emergence of new variants and the high mutation rate of SARS-CoV-2 challenge the efficacy of existing vaccines and have highlighted the need for novel vaccines that will be effective against SARS-CoV-2 variants. In this study, we exploit all four structural proteins of SARS-CoV-2 to execute a potential vaccine against SARS-CoV-2 and its variants. The vaccine was designed by utilizing the antigenic, non-toxic, and non-allergenic epitopes of B-cell and T-cell from conserved regions of viral structural proteins. To build a vaccine construct, epitopes were connected through different linkers and adjuvants to enhance the immunogenicity and specificity of the epitopes. The vaccine construct was selected through the aforementioned filters and it scored 0.6 against the threshold of 0.4 on VexiJen 2.0 which validates its antigenicity. Toll-like receptors (TLR2−4, and TLR8) and vaccine construct were docked by Cluspro 2.0, and TLR8 showed strong binding of -1577.1 kCal/mole. To assess the reliability of the docked complexes, C-IMMSIM's immune simulations over three doses of the vaccine and iMODS' molecular dynamic simulation were executed. The stability of the vaccine construct was evaluated through the physicochemical analyses and the findings suggested that the manufactured vaccine is stable under a wide range of circumstances and has the ability to trigger immune responses against various SARS-CoV-2 variants (due to conserved epitopes). However, in order to strengthen the vaccine formulation and assess its safety and effectiveness, additional studies and research are required to support the computational data of this research at In-vitro and In-vivo levels.

1. Introduction

An outbreak of pneumonia in Wuhan, China resulted in the emergence of COVID-19, a disease caused by the SARS-CoV-2 virus (Severe Acute Respiratory Syndrome Coronavirus 2). This extremely contagious disease quickly spread to 191 nations and affected over 75 million individuals, leading to more than 1.5 million fatalities. It had officially been classified worldwide as a pandemic and an international health crisis (Abebe et al., 2020). Victims of COVID-19 suffering from the viral infection caused by SARS-CoV2 experience a common cold along with fever-like symptoms similar to the flu; a small number of patients also develop a catastrophic type of pneumonia that inevitably leads to fatality (Boopathi et al., 2021; Nguyen et al., 2020; Vankadari et al., 2020). Furthermore, individuals with COVID-19 who ended up hospitalized in the intensive care unit (ICU) were reported to have abdominal pain, anorexia, cerebrovascular and cardiovascular disease (Chan et al., 2020; Wang et al., 2020). Infected individuals play a crucial role in spreading from person to person because they could pass on the virus to others via sneezing, coughing, exhales, and several other means of transmission (Bhattacharya et al., 2020; Chan et al., 2020).

The approximate 29.8−29.9 kb genome size of + ve sense single-stranded RNA (+ + ssRNA) of SARS-CoV-2 possesses a 5' end cap and poly-A tail at a 3' end. The sequence of its genome resembles those of the bat coronavirus and SARS-CoV, correspondingly, by 79.5% and 96%. Phylogenetical analysis reveals that it belongs to the Coronaviridae family, Nidovirales order, and member 2B group of β-coronavirus (Hui et al., 2020). Four structurally important proteins comprising SARS-CoV2 include spike (S), envelope (E), nucleocapsid (N), and membrane (M). Several vaccines were made available to reduce the severity and impact of the COVID-19 pandemic by mass vaccination. These clinically approved vaccines were based on killed/inactivated viruses, recombinant adenovirus, or mRNA of viral spike protein (Dong et al., 2020). The concept of an mRNA-based vaccination has been around for nearly 20 years, and recent evidence demonstrating its competence in both human and animal models against a variety of viral diseases, including Zika, HIV-1, rabies, and influenza virus, validates its efficacy (Alberer et al., 2017; Bahl et al., 2017; Richner et al., 2017). Up to this point, only two vaccines based on mRNA targeting SARS-CoV2 were successfully created: BNT162b2 (Pfizer) and mRNA-1273 (Moderna) (Baden et al., 2021; Polack et al., 2020).

SARS-CoV2 has been evolving over time, just like all other RNA viruses, as a result of random mutations in numerous viral genes, which produce a variety of genetic variations. The pressure of natural selection has sculpted this genetic variety, causing the emergence of viral variants with improved immune escape mechanisms, higher virulence, and enhanced transmission (Finkel et al., 2021). The aforementioned variations were designated as Variant of Interest (VOI) followed by Variant of Concern (VOC) to emphasize the extensive global surveillance and study. The global COVID-19 warning level is still elevated due to the emergence of new strains like Delta, Gamma, and Lambda (Bian et al., 2021; Martin et al., 2021). On the eleventh of November, 2021, the Omicron variation first emerged in Botswana. It subsequently spread to various other countries until it was confirmed in over 57 different countries globally (Petersen et al., 2022).

It is undeniable that the rapid progression of mutations will cause the SARS-COV-2 strains to quickly develop resistance to currently existing vaccinations, and that there will be an annual prerequisite for novel vaccines to protect against the fast-evolving flu viruses (Kotey et al., 2019). Therefore, this study attempted to design a single vaccine comprised of conserved immunogenic epitopes predicted from M-protein, E-protein, S-protein, and N-protein. This study also measured the physicochemical properties of B-cell and T-cell epitopes to characterize the vaccine and also screened for any allergic and toxic effects. The study also involved the molecular docking of various Toll-Like Receptors (TLRs) with the construct and stability of docked complexes ensured by MD simulations or Normal Mode Analysis (NMA). In the final step immune simulations were performed to confirm the antigenicity of the fabricated vaccine by estimating the production of various immune molecules and cells involved in both innate and adaptive immune responses.

2. Methodology

The supplementary file contains the access dates and links to all the tools mentioned.

1. Retrieval of Sequences Viral Proteins

The UniProt database was used to obtain the sequences of SARS-COV-2's structural proteins in the FASTA format. The acquired proteins comprised the M-protein, E-protein, S-protein, and N-protein (Satarker and Nampoothiri, 2020), each designated with distinct UniProt IDs - P0DTC5, P0DTC4, P0DTC2, and P0DTC9, correspondingly.

2.2. Predicting Antigenicity of viral proteins

In order to evaluate the antigenicity of the four selected proteins, they were screened on the VaxiJen v2.0 server as this server predicts the antigenicity of proteins by analyzing their physicochemical properties. The recommended threshold value of 0.4 was utilized to analyze the viral proteins, indicating that proteins with scores above this threshold are considered highly antigenic (Doytchinova and Flower, 2007; Naveed et al., 2022).

2.3. Predicting Physicochemical properties of viral proteins

To gain deeper insights into the physicochemical traits of the selected viral proteins, the ProtParam tool was used which is available on the ExPasy server (Gasteiger et al., 2005). This tool computes the proteins' multiple physical and chemical characteristics, which are essential for understanding the molecular and structural properties of the protein.

2.4. Prediction of B-cell restricted epitopes of viral proteins

In order to predict the linear B-cell epitopes of the chosen viral proteins, we used the B-cell Antigen Sequence Properties tool available on the IEDB server. This online tool employed the Bepipred Linear Epitope Prediction 2.0 method, which is designed to predict potential epitope sites in protein sequences. The tool utilized a threshold of 0.500 and a center position of 4, which are the recommended settings for this method. The adjusted threshold is compatible with the Random Forest algorithm that screens protein sequences to predict B-cell confined sequences (Jespersen et al., 2017). This algorithm works in two steps by first determining the results from crystal structures and then reporting the sequential predictions, which can be subjected to several filters for further analysis. This tool allowed the identification of potential linear B-cell epitopes within the viral proteins of SARS-COV-2.

2.5. T-Cells-Restricted Epitopes Prediction using IEDB Tools

In order to forecast the T-cell-specific epitopes, two different tools were utilized provided by IEDB. For predicting MHC-I binding epitopes, the recommended NetMHCpan EL 4.1 method was used (Reynisson et al., 2020), and for predicting MHC-II binding epitopes, IEDB Recommended 2.22 method was executed (Jensen et al., 2018). The same recommended sets of alleles and applied settings were used for both tools to ensure consistency throughout the study. As a result, we were able to predict potential T-cell-restricted epitopes from the chosen viral proteins.

2.6. Filtration and Profiling of Predicted Epitopes

The most optimal epitopes for the final vaccine design were selected through a rigorous selection process involving four filters. The first filter was the Vaxijen v2.0 tool for predicting antigenicity (Doytchinova and Flower, 2007), followed by the Toxinpred tool for toxicity prediction (Sharma et al., 2022). The third filter was the AllerTOP tool for allergenicity prediction (Dimitrov et al., 2013) and finally, the Align tool of the UniProt portal was applied to evaluate the conservancy of the epitopes (Coudert et al., 2022). Only those epitopes that passed through all four filters were selected to ensure the highest quality of the final vaccine design.

2.7. Percentage of filtered epitopes with population coverage

After identifying the most suitable epitopes for the vaccine design, the next step was to analyze the population coverage of each selected epitope. For this purpose, the researchers utilized the services of integrated population coverage tool available on the IEDB server (Bui et al., 2006). This tool allowed us to estimate the proportion of the world population that could potentially respond to each epitope by evaluating the interaction between the filtered epitopes and the maximum number of Major Histocompatibility complex (both type I and II) alleles provided to the tool.

2.8. Constructing the vaccine design

The vaccine was designed to include several components that work together to elicit an immune response against SARS-CoV-2. The design consisted of a 6X histidine marker, an adjuvant, B-cell-specific epitopes, and T-cell-specific epitopes that were connected in a series using specific linkers. To connect the selected epitopes, we used various linkers, including EAAK for B-cell epitopes, GPGPG for MHC-I epitopes, and AAY for MHC-II epitopes (Naveed et al., 2022). The vaccine construct was designed, to begin with the inclusion of the adjuvant Human Beta Defensin 3 with UniProt ID: Q5U7J2, which has been shown to enhance the immune response to vaccines (Ferris et al., 2013). The use of specific linkers to connect the selected epitopes is an important aspect of the vaccine design as these linkers ensured that the epitopes are presented to the immune system in a way that maximizes their effectiveness (Gokhale and Khosla, 2000; Parvizpour et al., 2020).

2.9. Profiling of Vaccine Construct

The newly developed vaccine underwent a thorough profiling process to evaluate its antigenic potential using the VaxiJen v2.0 tool, ensuring that the construct is enough antigenic to initiate a proper immune response against the viral particles (Doytchinova and Flower, 2007). After the vaccine construct was created, it underwent a series of analyses to ensure its safety and efficacy. First, the AllerTop v2.0 and ToxinPred tools were used to assess the vaccine for any

potential allergenicity and toxicity (Dimitrov et al., 2013; Sharma et al., 2022). These analyses are crucial in determining the safety of the vaccine and ensuring that it will not cause any harmful side effects in humans. Additionally, the physicochemical characteristics of the vaccine were analyzed using the ExPasy ProtParam server to evaluate the protein's stability, solubility, and other properties (Gasteiger et al., 2005). Finally, the vaccine underwent a secondary structure analysis using the PsiPred and SOPMA servers to predict protein folding and structural stability (Geourjon and Deléage, 1995; McGuffin et al., 2000). These analyses are necessary to ensure the vaccine is safe and enough effective to produce adequate immune responses against the SARS-COV-2 strains.

2.10. Three-Dimensional (3D) Modeling and Validations

The third-dimensional arrangement of the vaccine construct was determined by utilizing the I-TASSER server (Zhang, 2008). The scrutinized vaccine's 3D model undergoes a series of validations and refinements process using PROCHECK v6.0 and GalaxyRefine, respectively. PROCHECK generated a Ramachandran plot to evaluate the model, while GalaxyRefine was employed to modify the bond angles in the vaccine's 3D model (Heo et al., 2013; Wlodawer, 2017). To augment PROCHECK's validation outcomes, the ProSA-web server was also utilized (Wiederstein and Sippl, 2007).

2.11. Disulfide Engineering of Vaccine Construct

To evaluate the stability of the vaccine model, we employed the Disulfide tool on Design 2.0 (DbD2) (Craig and Dombkowski, 2013). This online server predicts potential disulfide linkages in the vaccine construct and presents the results in the form of a modified vaccine model (mutant model).

2.12. Molecular Docking with Toll-Like Receptors (TLRs)

The engagement of Toll-like receptors (TLRs) in the initiation of intracellular signaling pathways that culminate in the release of various immune compounds such as IRF3, NF-κB, inflammatory cytokines, and IL8 during the initial stages of viral infection has been well-established. To further explore this phenomenon, this study used the employed docking analysis of five TLRs, namely TLR2, TLR3, TLR4, TLR5, and TLR8, with 6NIG, 2A0Z, 4G8A, 3J0A, and 4R0A PDB IDs, respectively (Naveed et al., 2022). The Cluspro 2.0 server was utilized to carry out the docking (Kozakov et al., 2017) and the resulting docked complexes were then analyzed and visualized using the PyMol software.

2.13. Molecular Dynamics Simulations

The iMODS tool is founded upon Normal Mode Analysis (NMA), a valuable technique for determining the range of conformational states available to a given macromolecule and examining protein stability and mobility on a large scale. In this study, the iMODS server was utilized to render the docked complexes by modifying the force field over different time intervals (López-Blanco et al., 2014).

2.14. Codon optimization for virtual cloning

The process of codon optimization for the vaccine was initiated through the use of the Java Codon Adaptation Tool (JCAT) (Grote et al., 2005). This JCAT server algorithmically determines the most appropriate synonymous codons based on specific features of the target organism (E-coli K12), such as codon usage bias, GC content, and other factors. Once the nucleotide sequence was optimized, it was integrated into the pET28a (+) vector, which is commonly used for protein expression in E. coli, using SnapGene software (Drummond et al., 2022).

2.15. C-IMMSIM for Immune Simulation

The C-IMMSIM platform has been developed to anticipate and approximate the potential immune responses that may occur in the human body after the administration of the customized vaccine (Rapin et al., 2010). For this investigation, simulation steps were adjusted to 1050 in order to achieve the maximum elapsed time. The second modified parameter was the Time Steps of Injection for the second and third booster doses, which were adjusted to 84 and 170, respectively (Naveed et al., 2022).

3. Results

3.1. Physicochemical and Antigenicity Analysis of Viral Proteins:

In this study, an analysis was conducted on four candidate proteins to evaluate their antigenicity scores, namely E-protein, M-protein, N-protein, and S-protein. The VaxiJen algorithm was used to predict the antigenic nature of the proteins, and the obtained scores were 0.6025, 0.5102, 0.5059, and 0.4646, respectively. The proteins were considered antigenic as their scores exceeded the adjusted threshold. Notably, E-protein exhibited the highest antigenicity score while S-protein had the lowest score among the selected viral proteins. In addition, the physicochemical properties of the proteins were analyzed using ProtParam. The results showed that all four proteins were stable based on their predicted instability index (II). However, the GRAVY scores indicated that S-protein and N-protein were hydrophilic while E-protein and M-protein were hydrophobic. Further details of the physicochemical attributes are provided in Table 1.

Table 1
List of various physicochemical properties with Corresponding Uniprot IDs of selected viral proteins

Protein	UniProt ID	VaxiJen Score	Amino acids	Theoretical PI	Instability Index (II)	GRAVY	Aliphatic index
Spike	P0DTC2	0.4646	1273	6.24	33.01	-0.079	84.67
Envelope	P0DTC4	0.6025	75	8.57	38.68	1.128	144.00
Membrane	P0DTC5	0.5102	222	9.51	39.14	0.446	120.86
Nucleocapsid	P0DTC9	0.5059	419	10.07	55.09	-0.971	52.53

3.2. B-Cell Restricted Epitopes predictions:

The aim of the B cell epitope prediction was to identify potential antigens capable of efficiently interacting with B lymphocytes, as B-cell epitopes possess modified features that enable B-cells to recognize and activate potent immune responses against viral infections, thereby initiating the production of antibodies from plasma cells (a differentiated form of B-cells). After conducting the analysis, a total of 34 B-cell-specific epitopes were identified for the S-protein, 2 for the E-protein, 6 for the M-protein, and 11 for the N-protein. However, to ensure that only the most promising epitopes were selected for further analysis, we narrowed our focus to those with a length between 9–50 amino acids. To facilitate the analysis of these predicted B-cell-restricted epitopes, we have listed all of them in Supplementary Tables S1a, S1b, S1c, and S1d. Additionally, the graphs in Fig. 1 were generated using the IEDB tool, which allowed for the prediction of B-cell restricted epitopes based on threshold scores.

3.3. Predicting the T-Cell-Specific Epitopes

For the identification of potential epitopes that could be restricted to MHC-I and MHC-II of T-cells, the IEDB algorithm was employed. Numerous epitopes were identified for each viral protein, but we selected the top 30 based on their percentile scores. Three epitopes from the predicted sets of MHC-I and MHC-II epitopes were filtered out in the final for each viral protein. Using this approach, a total of 12 MHC-I restricted epitopes (three from each protein) and an equivalent number of MHC-II restricted epitopes were selected for the final vaccine construct. The epitopes were predicted based on the allele sets, which are provided in Supplementary Table S2a and S2b for both classes of MHCs (I & II), along with their respective peptide lengths.

3.4. Filtration and Profiling of Predicted Epitopes

In order to improve the efficacy of the final vaccine construct, a set of five filters were employed to screen the predicted epitopes, ensuring their antigenicity, non-toxicity, non-allergenicity, conservancy, and length ranging from 9–50 amino acids only for B-cell-restricted linear epitopes. Among the 34 predicted B-cell epitopes for the spike protein, only four met the filtering criteria, while only one out of two epitopes for the envelope protein passed. For the nucleoprotein, although 11 B-cell-restricted epitopes were identified, only three passed the filter tests and were chosen for further analysis and vaccine development. In contrast, all six B-cell-restricted linear epitopes predicted for the membrane protein were rejected by the filters, resulting in the exclusion of B-cell epitopes of this protein from the final vaccine construct. The final vaccine design included only the three epitopes with the most elevated antigenicity scores from each protein after filtering the top 30 MHC-I and MHC-II restricted T-cell epitopes. As a result, a total of 12 MHC-I restricted T-cell epitopes and 12 MHC-II restricted T-cell epitopes were included in the vaccine construct. Details of the epitope filtering process can be found in Tables 2a, 2b, and 2c, and Supplementary Figures S1a, S1b, S1c, and S1d provide the conservancy analysis.

Spike Glycoprotein

Start	End	Peptide	Length	Vaxijen Score	Antigenicity	Toxicity	Allergenicity	Conserv
206	221	KHTPINLVRDLPQGFS	16	0.6403	Antigen	Non-Toxin	Non-Allergen	Conserv
369	393	YNSASFSTFKCYGVSPTKLNDLCFT	25	1.4031	Antigen	Non-Toxin	Non-Allergen	Conserv
404	426	GDEVRQIAPGQTGKIADYNYKLP	23	1.1017	Antigen	Non-Toxin	Non-Allergen	Conserv
656	666	VNNSYECDIPI	11	0.6124	Antigen	Non-Toxin	Non-Allergen	Conserv

Envelope small membrane protein

Start	End	Peptide	Length	Vaxijen Score	Antigenicity	Toxicity	Allergenicity	Conserv
57	71	YVYSRVKNLNSSRVP	15	0.4492	Antigen	Non-Toxin	Non-Allergen	Conserv

Membrane protein

Start	End	Peptide	Length	Vaxijen Score	Antigenicity	Toxicity	Allergenicity	Conserv
N/A	N/A	-	-	-	-	-	-	-

Nucleoprotein

Start	End	Peptide	Length	Vaxijen Score	Antigenicity	Toxicity	Allergenicity	Conserv
59	105	HGKEDLKFPRGQGVPINTNSSPDDQIGYYRRATRRIRGGDGKMKDLS	47	0.5773	Antigen	Non-Toxin	Non-Allergen	Conserv
226	267	RLNQLESKMSGKGQQQQGQTVTKKSAAEASKKPRQKRTATKA	42	0.5627	Antigen	Non-Toxin	Non-Allergen	Conserv
276	299	RRGPEQTQGNFGDQELIRQGTDYK	24	0.6277	Antigen	Non-Toxin	Non-Allergen	Conserv

Table 2
b: The results of various filters that were applied in the selection of 12-MHC-I epitopes

Spike Glycoprotein

Allele	Start	End	Length	Peptide	IEDB Score	Vaxijen Score	Antigenicity	ToxinPred	AllerTOP	Conservancy
HLA-B*40:01	1016	1024	9	AEIRASANL	0.976223	0.7082	Antigen	Non-Toxin	Non-Allergen	Conserved
HLA-B*15:01	1264	1272	9	VLKGVKLHY	0.943479	1.2378	Antigen	Non-Toxin	Non-Allergen	Conserved
HLA-B*51:01	714	722	9	IPTNFTISV	0.942802	0.882	Antigen	Non-Toxin	Non-Allergen	Conserved

Envelope small membrane protein

Allele	Start	End	Length	Peptide	Score	Vaxijen Score	Antigenicity	ToxinPred	AllerTOP	Conservancy
HLA-A*31:01	61	69	9	RVKNLNSSR	0.939976	0.8998	Antigen	Non-Toxin	Non-Allergen	Conserved
HLA-A*68:01	30	38	9	TLAILTALR	0.682278	0.7223	Antigen	Non-Toxin	Non-Allergen	Conserved
HLA-A*30:02	49	57	9	VSLVKPSFY	0.652132	0.7476	Antigen	Non-Toxin	Non-Allergen	Conserved

Membrane protein

Allele	Start	End	Length	Peptide	Score	Vaxijen Score	Antigenicity	ToxinPred	AllerTOP	Conservancy
HLA-A*68:01	137	146	10	ELVIGAVILR	0.977462	0.9998	Antigen	Non-Toxin	Non-Allergen	Conserved
HLA-A*68:01	138	146	9	LVIGAVILR	0.948655	1.1027	Antigen	Non-Toxin	Non-Allergen	Conserved
HLA-B*57:01	84	92	9	MACLVGLMW	0.935424	0.7889	Antigen	Non-Toxin	Non-Allergen	Conserved

Nucleoprotein

Allele	Start	End	Length	Peptide	Score	Vaxijen Score	Antigenicity	ToxinPred	AllerTOP	Conservancy
HLA-A*03:01	361	370	10	KTFPPTEPKK	0.986352	0.7657	Antigen	Non-Toxin	Non-Allergen	Conserved
HLA-B*15:01	305	314	10	AQFAPSASAF	0.986204	0.5986	Antigen	Non-Toxin	Non-Allergen	Conserved
HLA-B*57:01	100	108	9	KMKDLSPRW	0.958352	1.7462	Antigen	Non-Toxin	Non-Allergen	Conserved

Spike Glycoprotein

Allele	Start	End	Length	Peptide	Vaxijen Score	Antigenicity	ToxinPred	AllerTOP	Conservancy
HLA-DRB1*13:02	115	129	15	QSLLIVNNATNVVIK	0.4343	Antigen	Non-Toxin	Non-Allergen	Conserved
HLA-DRB1*01:01	511	525	15	VVLSFELLHAPATVC	0.8618	Antigen	Non-Toxin	Non-Allergen	Conserved
HLA-DRB1*01:01	510	524	15	VVVLSFELLHAPATV	0.8083	Antigen	Non-Toxin	Non-Allergen	Conserved

Envelope small membrane protein

Allele	Start	End	Length	Peptide	Vaxijen Score	Antigenicity	ToxinPred	AllerTOP	Conservancy
HLA-DPA1*03:01/DPB1*04:02	18	32	15	LLFLAFVVFLLVTLA	0.8122	Antigen	Non-Toxin	Non-Allergen	Conserved
HLA-DPA1*03:01/DPB1*04:02	19	33	15	LFLAFVVFLLVTLAI	0.7471	Antigen	Non-Toxin	Non-Allergen	Conserved
HLA-DPA1*03:01/DPB1*04:02	21	35	15	LAFVVFLLVTLAILT	0.8229	Antigen	Non-Toxin	Non-Allergen	Conserved

Membrane protein

Allele	Start	End	Length	Peptide	Vaxijen Score	Antigenicity	ToxinPred	AllerTOP	Conservancy
HLA-DQA1*01:01/DQB1*05:01	49	63	15	IKLIFLWLLWPVTLA	0.8704	Antigen	Non-Toxin	Non-Allergen	Conserved
HLA-DPA1*01:03/DPB1*02:01	88	102	15	VGLMWLSYFIASFRL	0.6658	Antigen	Non-Toxin	Non-Allergen	Conserved
HLA-DQA1*01:01/DQB1*05:01	50	64	15	KLIFLWLLWPVTLAC	0.7344	Antigen	Non-Toxin	Non-Allergen	Conserved

Nucleoprotein

Allele	Start	End	Length	Peptide	Vaxijen Score	Antigenicity	ToxinPred	AllerTOP	Conservancy
HLA-DRB1*09:01	305	319	15	AQFAPSASAFFGMSR	0.5266	Antigen	Non-Toxin	Non-Allergen	Conserved
HLA-DRB1*11:01	84	98	15	IGYYRRATRRIRGGD	0.6649	Antigen	Non-Toxin	Non-Allergen	Conserved
HLA-DRB1*07:01	328	342	15	GTWLTYTGAIKLDDK	0.9934	Antigen	Non-Toxin	Non-Allergen	Conserved

3.5. Epitopes' Population-Coverage Analysis

The combined population coverage of the 12 MHC-I restricted epitopes selected for the vaccine construct was found to be 98.55% for the global population. Similarly, for the 12 MHC-II restricted epitopes, the coverage was determined to be 72.18%. The population coverage scores were determined based on the selected alleles used in the analysis, which are provided in Supplementary Table S3a and S3b for individual epitopes. The cumulative population coverage for epitopes restricted to both MHC classes exceeded 50%, which indicates the potential effectiveness of the vaccine construct. The Fig. 2 illustrates the individual epitope coverage for both MHC classes. Additional information on individual epitope coverage scores and the percentage genotype frequencies of HLA alleles can be found in Supplementary Tables S4a, S4b, S4c, and S4d.

3.6. Vaccine construct's assembly:

The vaccine components were systematically assembled in sequential order, commencing with the adjuvant sequence attached to the 6xHis tag, followed by the B-cell-restricted epitopes joined together with EAAK linkers. The MHC-I and MHC-II epitopes selected earlier were then attached to the chain using GPGPG and AYY linkers, respectively. To complete the chain, another 6x His tag was added. The inclusion of these tags, one at the Amino (N) terminus and the other at the Carbon (C) terminus serves the purpose of identifying and purifying the cloned protein both in-vitro and in-vivo. Figure 3 shows a detailed representation of the vaccine construct.

3.7. Quality Check and Physicochemical properties

The VexiJen 2.0 predicted that the construct had a 0.6019 score, which validated the antigenicity of the vaccine, similarly, the ToxinPred and AllerTop also showed that the construct is non-toxin and non-allergen in nature, respectively. These three tools validated the quality and safety of the vaccine as it can

generate proper immune responses without the plausibility of side effects (i.e., toxicity and hypersensitivity or allergic reactions). The ProtParam results in Table 3 showed that the vaccine construct constituted 687 amino acids with a molecular weight of 73725.68 Da. The Instability Index (II) score of 34.83 substantiated the stability of the construct in the test tube. The 87.87 computed Aliphatic Index score of the construct validated that the protein is thermostable which makes it a viable candidate to initiate an immunogenic response without any changes in the structural confirmation. As the immune response begins the body temperature changes due to the inflammation but the construct can withstand such changes without losing the structural integrity. The sturdy structural folding of the vaccine was corroborated through the theoretical PI of 9.90 and it will be helpful in the implementation of purification methods. The constructed vaccine is hydrophilic in nature as it has negative Grand Average of Hydropathicity (GRAVY) scores (-0.028). This hydrophilicity of the construct will promote the immunogenic effects as it allowed the interaction of the construct with the intracellular and extracellular water residues of the host. The construct has a short half-life of 3.5 hours and 30 hours with and without 6x His tag respectively in mammalian reticulocytes (in-vitro). This short half-life span illustrated that the antigens of the construct remained in the host only to develop immunity without acting as the pathogen. It also has a half-life of > 10 hours in Escherichia coli (In-vivo) allowing enough time for the production and extraction of the vaccine during the cloning process.

Table 3
Results of various physicochemical properties of vaccine construct along with VaxiJen score

Protein	VaxiJen Score	Amino acids	Molecular Weight (Da)	Theoretical PI	Instability Index (II)	GRAVY	Aliphatic index	Half-life (with 6x His tag)	Half-life (without 6x His tag)
Vaccine	0.6019	687	73725.68	9.90	34.83	-0.028	87.87	3.5 hours	30 hours

3.8. Secondary Structure Profiling

The secondary structure composition of the construct was analyzed using the SOPMA server, which revealed that the protein contained 264 residues (38.43%) of alpha helix, 127 residues (18.49%) of extended strands, and 296 residues (43.09%) of coils. The results obtained from the SOPMA server were validated by comparing them with the results from PsiPred, as presented in Fig. 4. The trans-membrane region of the construct was identified to consist of residues 483 to 632, with residues 483 to 513 comprising the pore-lining of the plasma membrane of the host cells, according to the MEMSAT-SVM schematic of PsiPred. The MEMSAT-SVM schematic of the construct is also depicted in Fig. 4.

3.9. Three-Dimensional (3D) Modeling and Validations

Five distinct three-dimensional (3D) models of the vaccine construct were produced using I-TASSAR, with the one having the highest C-score of -2.08 being selected for further analysis. The chosen model had TM-score and RMSD values of 0.47 ± 0.15 and 13.1 ± 4.1Å, respectively, while Supplementary Table S5 contains the C-scores of all the predicted 3D models. The model was evaluated by PROCHECK for Ramachandran analysis, which revealed that the initial model had only 78.3% of the residues in the most favored region, which was not sufficient for the validation of the 3D model. Therefore, the GalaxyRefine server was applied to refine the selected model, leading to a refined model where 90.1% of the residues were situated in the most favored region, 7.3% of the residues in the additional allowed region, 1.1% of the residues in the generously allowed region, and 1.6% of the residues in the disallowed region. The Ramachandran analysis of the refined model with 90.1% of the residues in the most favored region validated the accuracy of the model. This was further supported by the ProSA-web evaluation, which predicted a -4.8 Z-score for the model. The 3D structures of the raw and refined models are provided in Fig. 5, and the Ramachandran plot and ProSA-web evaluation charts are presented in Fig. 6.

3.10. Disulfide Engineering for stability

By utilizing the DbD2 server, the tertiary structure of the construct was analyzed, and it identified 54 residue pairs that potentially can form disulfide linkages. Out of these pairs, only seven were considered appropriate for disulfide engineering, as they exhibited maximum binding energy of up to 2.20 kcal/mol. This selection was based on the fact that 2.20 kcal/mol is the minimum energy required for the formation of a disulfide bond. The DbD2 generated a mutated wireframe tertiary model of the construct with the disulfide bonds between the selected residues shown in yellow color in Supplementary Figure S2. The original wireframe structure was also included in Supplementary Figure S2, while Supplementary Table S6 provided a list of all predicted and selected residue pairs for disulfide engineering.

3.11. Toll-Like Receptors (TLRs) Docked with Vaccine:

The underlying logic for conducting molecular docking between the vaccine and TLRs is based on the fact that these receptors play a vital role as the first-line immunogenic defense system for the activation of the innate immune response against viral infections. Figure 7 illustrates the selected docked complexes of TLRs with the vaccine model. For instance, Fig. 7A is depicting the complex of the TLR2 + vaccine model with the binding energy of -1107.7 kCal/mole and 46 cluster members. The docked complex of TLR3 and vaccine model has – 1146.7 kCal/mole binding energy with 44 cluster members (Fig. 7B), and the docking of TLR4 + vaccine model involved 33 cluster members with the binding energy of -1235.2 kCal/mole (Fig. 7C). Similarly, the TLR5 + vaccine complex represented in Fig. 7D has a binding energy of -1497.2 kCal/mole with 42 cluster members while the binding energy of the TLR8 + vaccine complex is -1577.1 kCal/mole with 77 cluster members is given in Fig. 7E. The models were selected based on two factors the first one is the highest negative binding energy as it indicates the stabilization of the reaction through tremendous binding affinities and mends the path toward the product. However, the second factor is the highest cluster members, which strengthen the intensity of binding affinities of the docked complex. The calculations in this research insinuated that the TLR8 has the highest negative binding energy and cluster members, enabling the receptor to become a predominant component of innate immune response followed by the activation of adaptive immune responses in case of infection. The remaining TLRs also had the tendency to generate an appropriate immune response upon interacting with the non-self-epitopes. Supplementary Table S7 is demonstrating the lowest energies, center values, and cluster members of docked complexes obtained from the analysis by the ClusPro 2.0 server.

3.12. Molecular Dynamic (MD) Simulations

The stability and mobility of the docked complexes were determined by molecular dynamics (MD) simulations, which investigated the conformational states of the macromolecules in a virtual biomolecular system upon perturbation. To render the docked complex, the iMODS server adjusted the force field at various time intervals and calculated several properties including individual and cumulative variance, deformability, eigenvalues, B-factor, covariance map, and elastic network depicted in Fig. 8 (TLR8 + vaccine). Figure 8A shows the deformability of the docked complex at individual residue levels, which is dependent on the distortion of C-α atoms. The average root-mean-square fluctuation (RMSF) of each residue, calculated in correspondence to the NMA mobility factor 8π2, is presented in the B-factor value graph of the stable docked complex in Fig. 8B. The eigenvalues, which demonstrate the motion stiffness of the complex and are directly linked with the energy required for structure deformation, are shown in Fig. 8C. The binding interaction between the protein and the ligand was stabilized and made flexible by a low eigenvalue of 6.6173e – 06. In Fig. 8D, the eigenvalue and variance, which are inversely related, are illustrated using red and green colors to represent them individually and cumulatively. Using the C-α Cartesian coordinates, the covariance map was generated to demonstrate the concatenation between the residues. Figure 8E visually presents the motion analysis of coupled residues by using blue, red, and white colors to represent anti-correlated, correlated, and uncorrelated motions, respectively. The spring network model delineated in Fig. 8F demonstrates that the Cα atoms in the docked molecules are connected by springs of varying strengths, represented by dots. The strength of each dot (spring) is determined by the darkness of the color in the graph, as lighter grays indicate less stiffness while darker colors represent the higher stiffness of the springs. The iMODS graphs of the remaining complexes are provided in Supplementary Figures S3a, S3b, S3c, and S3d respectively, while Supplementary Table S8 presents the eigenvalue of all the docked complexes.

3.13. Codon Adaptation and In-Silico Cloning

The vaccine construct's codon was optimized for expression in the E. coli K-12 strain, resulting in a refined series of 2061 nucleotides. The optimization process ensured that the Codon Adaptation Index (CAI) was 0.97, and the GC content was 54.1%, both of which are within the optimal range, thus increasing the probability of protein expression. In Fig. 9a, the graphs is illustrating the optimized CAI values, while the optimized nucleotide sequence is shown in Supplementary Figure S4. Adaptors in the form of EcoR1 (GAATTC) and BamH1 (GGATCC) restricted endonucleases were added to the beginning and end of the optimized nucleotide sequence, respectively, to facilitate its integration into the pET28a (+) plasmid. In-silico cloning was carried out using SnapGene software, resulting in a cloned recombinant plasmid with a length of 7.438 kb illustrated in Fig. 9b.

3.14. Immune Simulation by C-IMMSIM

The C-IMMSIM algorithm computed the host's immune responses upon the administration of the vaccine. The simulation predicted elapsed time of 280 days against 1050 simulation steps for 3 doses at 4-week intervals. Figure 10a presents the significant production of cytokines, and danger (D) levels (depicted in a sub-graph Fig. 10a). The adaptive immune response is essential for the development of long-term immunity which depends on the T-cell and B-cell generation upon the interaction with the vaccine and Fig. 10 and Fig. 11 described their surge. However, Figs. 11a and 11b demonstrate the memory B-cell population and B-cell population per state (cells/mm3) respectively, while Figs. 12a and 12b illustrate the sudden rise of Helper T-cells (TH-cells) and cytotoxic T-cells (TC-cells), correspondingly. The vaccine also exhibited significant efficacy in producing various types of antibodies, including IgM, IgA, IgG1 + IgG2, IgM + IgG, IgG1, and IgG2, as presented in Fig. 10b. Supplementary Figures S5 and S6 provide the production levels of other immune cells.

4. Discussion

Vaccines have played a crucial role in the history of public health, providing an effective means of immunization against newly emerging pathogens. However, traditional methods of vaccine design often involve the use of entire organisms or large proteins, which can trigger adverse and unwanted hypersensitive reactions (Jorge and Dellagostin, 2017). With the progress in bioinformatics, the process of vaccine design has also undergone remarkable advancements, making it a much faster and more efficient process. In recent times, bioinformatics tools have been great contributors to the identification and analysis of potential epitopes associated with different pathogens, leading to the development of novel vaccine designs and strategies (Arya and Bhatt, 2021). So, in this study, we employed a promising field of immunoinformatics to create a multi-epitope vaccine that targets various structural proteins of SARS-CoV-2, the pathogen behind the COVID-19 pandemic. To effectuate this, a combination of computational tools and algorithms was employed to detect the immunogenic/antigenic epitopes in the viral proteins and subsequently create a chimeric peptide that could generate a dynamic and precise immune reaction against SARS-CoV-2 while reducing the possibilities of adverse effects.

Since the emergence of SARS-CoV-2, numerous studies have been conducted to develop effective vaccines against the virus. Several of these studies have focused on developing multi-epitope vaccines that target either the S-protein (Kar et al., 2020), M-protein (Bashir et al., 2021), E-protein (Adam, 2021), or N-protein (Kumar et al., 2021; Oliveira et al., 2020). This study had a different approach compared to others that aimed to develop multi-epitope vaccines focusing on specific proteins of the SARS-CoV-2. Instead, the study designed a vaccine that would target all the virus's structural proteins to achieve a comprehensive immunization against the viral infection. The vaccine construct is composed of 12 MHC class-I epitopes specified to cytotoxic T-leukocytes (CTLs), 12 MHC class-II epitopes specified/restricted to Helper T-leukocytes (HTLs), and 8 linear B-cell-specified epitopes, which were selected carefully from each protein through the rigorous barrier of filters that ensured the occurrence of both adaptive and innate immune responses. The goal was to activate the immune system by targeting all the structural proteins, resulting in production of a diverse range of epitope-specific antibodies against different various SARS-COV-2 strains. Notably, the inclusion of MHC-I and MHC-II restricted T-cell epitopes plays a crucial role in activating cell-mediated immune responses, which can target infected cells and eliminate viral particles (Broere and van Eden, 2019). In addition, the presence of linear epitopes restricted to B-cells enhances the humoral immune response by inducing the production of neutralizing antibodies (Cancro and Tomayko, 2021).

Creating effective vaccines against viruses has always been a major hurdle, primarily due to the rapid mutation rates displayed by viruses. As viruses evolve, they tend to develop new variants that can evade the immune response generated by the existing vaccines, thereby rendering them less effective (Wang et al., 2021). The characteristic of viruses to mutate rapidly makes them resistant to vaccines through antigenic shift or drift, thus necessitating the need for new vaccines each year against different viral infections such as swine flu (H1N1) which require refurbished vaccine almost every year (Bellino et al., 2019; Treanor, 2004). Similarly, the SARS-CoV-2 virus has existed in several variants since the outbreak of the pandemic, posing a significant challenge to vaccine development efforts (Harvey et al., 2021). In this research, we aimed to address this challenge by selecting epitopes from the conserved regions of the viral proteins. This approach ensured the longevity of the vaccine construct, as the conserved regions are less prone to mutation and reduced the risk of resistance against the vaccine. The final vaccine construct developed in this study is a hybrid of conserved B-cell and T-cell epitopes from all the structural proteins. This design aimed to generate a cascade of immune responses while ensuring the safety of the vaccine, as it is non-allergen, non-toxin, and has a short half-life. The use of conserved epitopes in the vaccine design also ensures that the vaccine can provide protection against multiple variants of the virus which can reduce the need for frequent updates or refurbishments in the vaccine over time.

The efficacy and safety of epitope-based vaccines rely on several crucial factors, such as their ability to elicit a robust immune response while remaining safe for the recipient (Dutta and Langenburg, 2023). With these considerations in mind, the selected epitopes for this vaccine experienced a rigorous screening process to ensure the quality and safety of the final construct. Existing vaccines against SARS-COV-2 have reported various side effects, including anaphylaxis, mild headache, tiredness, and skin rashes. Therefore, each epitope included in the final construct was screened for any potential allergenic and toxic reaction in the living systems. Only those epitopes that passed all criteria were selected for inclusion in the final construct of the vaccine, while those that did not meet the required standards were completely rejected. This careful selection process helped us to minimize the risk of adverse reactions in vaccinated individuals and promote the safety and effectiveness of the vaccine.

The stability of the vaccine construct is a critical factor in ensuring its efficacy and longevity in the patient's body (Kim et al., 2022). Several physicochemical analyses were performed to assess the construct's stability, indicating that the protein is inherently stable (with instability index = 34.83). The high aliphatic index (87.87) suggests that it can withstand temperature fluctuations in the body of the host upon inflammatory reaction. Similarly, disulfide engineering was also employed to enhance and maintain the stability of tertiary structure with at least 7 identified disulfide linkages in the mutant model. In addition, 3-D modeling was used to validate the structure of the vaccine construct, with the Ramachandran plot demonstrating that over 90% of the residues were in the most favorable region and only a small percentage of residues (1.6%) were found in the disallowed region.

In the case of viral infections, TLRs are members of the first-line immunogenic defense system that detect the conserved pathogen-associated molecular patterns (PAMP) to initiate the immune response through a wide range of signaling pathways (Kaur et al., 2022). This research involved analyzing the binding energies and cluster members of various docked complexes between TLRs and the vaccine model. The results indicated that TLR8 had the highest negative binding energy and cluster members. Moreover, this study also utilized the normal mode analysis (NMA) of the iMODS server to examine the stability and dynamic properties of the docked complex formed by our designed vaccine and TLRs. By analyzing the data, we observed that the proteins in the complex showed minimal deformation at each residue. Additionally, eigenvalues were calculated for the docked complex (vaccine + TLR8), which exhibited a value of $6.6173e-06$, thus validating the reliability of our in-silico model for the vaccine. Such a comprehensive NMA analysis is a crucial step for the development of an effective vaccine candidate, as it can ensure the stability and functionality of the model within the dynamics of the host system in real-time (Kirar et al., 2022).

The primary objectives of vaccines are to generate a vigorous immune response and provide durable immunity against specific viral infections (Moss, 2022). In this study, the antigenicity of the vaccine construct was determined by the VaxiJen score of 0.6019, which exceeded the threshold of 0.4 indicating the antigenic nature of the vaccine construct. Nevertheless, to further enhance the potency and longevity of the immune response against the viral particles, an adjuvant, Human Beta Defensin 3, was incorporated at the beginning of the construct to magnify the effects of selected epitopes. As we know the adjuvants are known to modulate the immunogenicity of epitopes while reducing the need for a large number of antigens to induce the same response (Naveed et al., 2022). The antigenicity results were confirmed by immune simulations, which demonstrated the generation of immunogenic cells and molecules that are required for the establishment of long-term immunity against various SARS-COV-2 variants.

Study Limitation

As this study is based on computational algorithms and In-silico data, which means it may not accurately reflect the real-world experimental validations, so, it is essential to validate the algorithmic prediction through In-vitro and In-vivo experimentations. Secondly, even though there are advancements in computational technologies but still risks are there that some important epitopes may have been missed in the process due to the limitation of tools to understand the dynamic nature of pathogens. The binding affinity between an epitope and the major histocompatibility complex (MHC) molecule is a critical step in antigen presentation and T-cell activation. So, accurate predictions about this binding affinity are a big challenge, which may impact the efficacy of the vaccine construct designed in this study.

5. Conclusion

This study demonstrates that the designed vaccine has the potential to overcome the perpetual damage of SARS-CoV2 and its variants. Our computationally designed hybrid vaccine has epitopes from conserved regions of all structural proteins which makes it effective against the emerging strains of SARS-CoV-2. The results elicit that this vaccine tends to stimulate a powerful immune response and provide persistent immunization against the SARS-CoV-2 infection. We concur that our vaccine can impede COVID-19 from igniting unending havoc. Despite efforts that were previously made to prevent the spread of the pandemic, the appearance of new strains emphasizes the continuous threat of COVID-19 and demands new approaches to tackle this threat. The world has witnessed tremendous pain and disruption during the pandemic but we anticipate that our findings will provide the groundwork for future study in this area, and

eventually contribute to better world health accomplishments. However, the work is based on computational analyses which required in-vitro and in-vivo data for confirmational application of the vaccine construct in real-world settings.

Declarations

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: More data related to this study can be accessed by sending a rea-sonable email to 123allah.rakha@gmail.com.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Abebe, E.C., T.A. Dejenie, M.Y. Shiferaw, and T.J.V.j. Malik. 2020. The newly emerged COVID-19 disease: a systemic review. 17:1-8.

2. Adam, K.M. 2021. Immunoinformatics approach for multi-epitope vaccine design against structural proteins and ORF1a polyprotein of severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2). Tropical Diseases, Travel Medicine and Vaccines. 7:22.

3. Alberer, M., U. Gnad-Vogt, H.S. Hong, K.T. Mehr, L. Backert, G. Finak, R. Gottardo, M.A. Bica, A. Garofano, S.D. Koch, M. Fotin-Mleczek, I. Hoerr, R. Clemens, and F. von Sonnenburg. 2017. Safety and immunogenicity of a mRNA rabies vaccine in healthy adults: an open-label, non-randomised, prospective, first-in-human phase 1 clinical trial. The Lancet. 390:1511-1520.

4. Arya, H., and T.K. Bhatt. 2021. Chapter 20 - Role of Bioinformatics in Subunit Vaccine Design. In Molecular Docking for Computer-Aided Drug Design. M.S. Coumar, editor. Academic Press. 425-439.

5. Baden, L.R., H.M. El Sahly, B. Essink, K. Kotloff, S. Frey, R. Novak, D. Diemert, S.A. Spector, N. Rouphael, C.B. Creech, J. McGettigan, S. Khetan, N. Segall, J. Solis, A. Brosz, C. Fierro, H. Schwartz, K. Neuzil, L. Corey, P. Gilbert, H. Janes, D. Follmann, M. Marovich, J. Mascola, L. Polakowski, J. Ledgerwood, B.S. Graham, H. Bennett, R. Pajon, C. Knightly, B. Leav, W. Deng, H. Zhou, S. Han, M. Ivarsson, J. Miller, and T. Zaks. 2021. Efficacy and Safety of the mRNA-1273 SARS-CoV-2 Vaccine. The New England journal of medicine. 384:403-416.

6. Bahl, K., J.J. Senn, O. Yuzhakov, A. Bulychev, L.A. Brito, K.J. Hassett, M.E. Laska, M. Smith, Ö. Almarsson, J. Thompson, A. Ribeiro, M. Watson, T. Zaks, and G. Ciaramella. 2017. Preclinical and Clinical Demonstration of Immunogenicity by mRNA Vaccines against H10N8 and H7N9 Influenza Viruses. Molecular Therapy. 25:1316-1327.

7. Bashir, Z., S.U. Ahmad, B.H. Kiani, Z. Jan, N. Khan, U. Khan, I. Haq, F. Zahir, A. Qadus, and T. Mahmood. 2021. Immunoinformatics approaches to explore B and T cell epitope-based vaccine designing for SARS-CoV-2 Virus. Pakistan journal of pharmaceutical sciences. 34:345-352.

8. Bellino, S., A. Bella, S. Puzelli, A. Di Martino, M. Facchini, O. Punzo, P. Pezzotti, M.R. Castrucci, and G. the InfluNet Study. 2019. Moderate influenza vaccine effectiveness against A(H1N1)pdm09 virus, and low effectiveness against A(H3N2) subtype, 2018/19 season in Italy. Expert Review of Vaccines. 18:1201-1209.

9. Bhattacharya, M., A.R. Sharma, P. Patra, P. Ghosh, G. Sharma, B.C. Patra, S.S. Lee, and C. Chakraborty. 2020. Development of epitope-based peptide vaccine against novel coronavirus 2019 (SARS-COV-2): Immunoinformatics approach. Journal of medical virology. 92:618-631.

10. Bian, L., Q. Gao, F. Gao, Q. Wang, Q. He, X. Wu, Q. Mao, M. Xu, and Z. Liang. 2021. Impact of the Delta variant on vaccine efficacy and response strategies. Expert Review of Vaccines. 20:1201-1209.

11. Boopathi, S., A.B. Poma, P.J.J.o.B.S. Kolandaivel, and Dynamics. 2021. Novel 2019 coronavirus structure, mechanism of action, antiviral drug promises and rule out against its treatment. 39:3409-3418.

12. Broere, F., and W. van Eden. 2019. T Cell Subsets and T Cell-Mediated Immunity. In Nijkamp and Parnham's Principles of Immunopharmacology. M.J. Parnham, F.P. Nijkamp, and A.G. Rossi, editors. Springer International Publishing, Cham. 23-35.

13. Bui, H.H., J. Sidney, K. Dinh, S. Southwood, M.J. Newman, and A. Sette. 2006. Predicting population coverage of T-cell epitope-based diagnostics and vaccines. BMC Bioinformatics. 7:153.

14. Cancro, M.P., and M.M. Tomayko. 2021. Memory B cells and plasma cells: The differentiative continuum of humoral immunity. 303:72-82.

15. Chan, J.F., S. Yuan, K.H. Kok, K.K. To, H. Chu, J. Yang, F. Xing, J. Liu, C.C. Yip, R.W. Poon, H.W. Tsoi, S.K. Lo, K.H. Chan, V.K. Poon, W.M. Chan, J.D. Ip, J.P. Cai, V.C. Cheng, H. Chen, C.K. Hui, and K.Y. Yuen. 2020. A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. Lancet (London, England). 395:514-523.

16. Coudert, E., S. Gehant, E. de Castro, M. Pozzato, D. Baratin, T. Neto, C.J.A. Sigrist, N. Redaschi, A. Bridge, and T.U. Consortium. 2022. Annotation of biologically relevant ligands in UniProtKB using ChEBI. Bioinformatics. 39.

17. Craig, D.B., and A.A. Dombkowski. 2013. Disulfide by Design 2.0: a web-based tool for disulfide engineering in proteins. BMC Bioinformatics. 14:346.

18. Dimitrov, I., D.R. Flower, and I. Doytchinova. 2013. AllerTOP - a server for in silico prediction of allergens. BMC Bioinformatics. 14:S4.

19. Dong, Y., T. Dai, Y. Wei, L. Zhang, M. Zheng, and F. Zhou. 2020. A systematic review of SARS-CoV-2 vaccine candidates. Signal Transduction and Targeted Therapy. 5:237.

20. Doytchinova, I.A., and D.R. Flower. 2007. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. BMC Bioinformatics. 8:4.

21. Drummond, E., T. Kavanagh, G. Pires, M. Marta-Ariza, E. Kanshin, S. Nayak, A. Faustin, V. Berdah, B. Ueberheide, and T. Wisniewski. 2022. The amyloid plaque proteome in early onset Alzheimer's disease and Down syndrome. Acta Neuropathologica Communications. 10:53.

22. Dutta, S.K., and T. Langenburg. 2023. A Perspective on Current Flavivirus Vaccine Development: A Brief Review. In Viruses. Vol. 15.

23. Ferris, L.K., Y.K. Mburu, A.R. Mathers, E.R. Fluharty, A.T. Larregina, R.L. Ferris, and L.D. Falo, Jr. 2013. Human beta-defensin 3 induces maturation of human langerhans cell-like dendritic cells: an antimicrobial peptide that functions as an endogenous adjuvant. The Journal of investigative dermatology. 133:460-468.

24. Finkel, Y., O. Mizrahi, A. Nachshon, S. Weingarten-Gabbay, D. Morgenstern, Y. Yahalom-Ronen, H. Tamir, H. Achdout, D. Stein, O. Israeli, A. Beth-Din, S. Melamed, S. Weiss, T. Israely, N. Paran, M. Schwartz, and N. Stern-Ginossar. 2021. The coding capacity of SARS-CoV-2. Nature. 589:125-130.

25. Gasteiger, E., C. Hoogland, A. Gattiker, S.e. Duvaud, M.R. Wilkins, R.D. Appel, and A. Bairoch. 2005. Protein Identification and Analysis Tools on the ExPASy Server. In The Proteomics Protocols Handbook. J.M. Walker, editor. Humana Press, Totowa, NJ. 571-607.

26. Geourjon, C., and G. Deléage. 1995. SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. Computer applications in the biosciences : CABIOS. 11:681-684.

27. Gokhale, R.S., and C.J.C.o.i.c.b. Khosla. 2000. Role of linkers in communication between protein modules. 4:22-27.

28. Grote, A., K. Hiller, M. Scheer, R. Münch, B. Nörtemann, D.C. Hempel, and D. Jahn. 2005. JCat: a novel tool to adapt codon usage of a target gene to its potential expression host. Nucleic acids research. 33:W526-531.

29. Harvey, W.T., A.M. Carabelli, B. Jackson, R.K. Gupta, E.C. Thomson, E.M. Harrison, C. Ludden, R. Reeve, A. Rambaut, S.J. Peacock, D.L. Robertson, and C.-G.U. Consortium. 2021. SARS-CoV-2 variants, spike mutations and immune escape. Nature Reviews Microbiology. 19:409-424.

30. Heo, L., H. Park, and C. Seok. 2013. GalaxyRefine: Protein structure refinement driven by side-chain repacking. Nucleic acids research. 41:W384-388.

31. Hui, D.S., I.A. E, T.A. Madani, F. Ntoumi, R. Kock, O. Dar, G. Ippolito, T.D. McHugh, Z.A. Memish, C. Drosten, A. Zumla, and E. Petersen. 2020. The continuing 2019-nCoV epidemic threat of novel coronaviruses to global health - The latest 2019 novel coronavirus outbreak in Wuhan, China. International journal of infectious diseases : IJID : official publication of the International Society for Infectious Diseases. 91:264-266.

32. Jensen, K.K., M. Andreatta, P. Marcatili, S. Buus, J.A. Greenbaum, Z. Yan, A. Sette, B. Peters, and M. Nielsen. 2018. Improved methods for predicting peptide binding affinity to MHC class II molecules. Immunology. 154:394-406.

33. Jespersen, M.C., B. Peters, M. Nielsen, and P. Marcatili. 2017. BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. Nucleic acids research. 45:W24-w29.

34. Jorge, S., and O.A. Dellagostin. 2017. The development of veterinary vaccines: a review of traditional methods and modern biotechnology approaches. Biotechnology Research and Innovation. 1:6-13.

35. Kar, T., U. Narsaria, S. Basak, D. Deb, F. Castiglione, D.M. Mueller, and A.P. Srivastava. 2020. A candidate multi-epitope vaccine against SARS-CoV-2. Scientific Reports. 10:10895.

36. Kaur, A., J. Baldwin, D. Brar, D.B. Salunke, and N. Petrovsky. 2022. Toll-like receptor (TLR) agonists as a driving force behind next-generation vaccine adjuvants and cancer therapeutics. Current Opinion in Chemical Biology. 70:102172.

37. Kim, S.C., S.S. Sekhon, W.-R. Shin, G. Ahn, B.-K. Cho, J.-Y. Ahn, and Y.-H. Kim. 2022. Modifications of mRNA vaccine structural elements for improving mRNA stability and translation efficiency. Molecular & Cellular Toxicology. 18:1-8.

38. Kirar, M., H. Singh, and N. Sehrawat. 2022. Virtual screening and molecular dynamics simulation study of plant protease inhibitors against SARS-CoV-2 envelope protein. Informatics in Medicine Unlocked. 30:100909.

39. Kotey, E., D. Lukosaityte, O. Quaye, W. Ampofo, G. Awandare, and M. Iqbal. 2019. Current and Novel Approaches in Influenza Management. In Vaccines. Vol. 7.

40. Kozakov, D., D.R. Hall, B. Xia, K.A. Porter, D. Padhorny, C. Yueh, D. Beglov, and S. Vajda. 2017. The ClusPro web server for protein-protein docking. Nature protocols. 12:255-278.

41. Kumar, J., R. Qureshi, S.R. Sagurthi, and I.A. Qureshi. 2021. Designing of Nucleocapsid Protein Based Novel Multi-epitope Vaccine Against SARS-COV-2 Using Immunoinformatics Approach. International Journal of Peptide Research and Therapeutics. 27:941-956.

42. López-Blanco, J.R., J.I. Aliaga, E.S. Quintana-Ortí, and P. Chacón. 2014. iMODS: internal coordinates normal mode analysis server. Nucleic acids research. 42:W271-276.

43. Martin, D.P., S. Weaver, H. Tegally, E.J. San, S.D. Shank, E. Wilkinson, A.G. Lucaci, J. Giandhari, S. Naidoo, Y. Pillay, L. Singh, R.J. Lessells, R.K. Gupta, J.O. Wertheim, A. Nekturenko, B. Murrell, G.W. Harkins, P. Lemey, O.A. MacLean, D.L. Robertson, T. de Oliveira, and S.L. Kosakovsky Pond. 2021. The emergence and ongoing convergent evolution of the N501Y lineages coincides with a major global shift in the SARS-CoV-2 selective landscape. medRxiv : the preprint server for health sciences.

44. McGuffin, L.J., K. Bryson, and D.T. Jones. 2000. The PSIPRED protein structure prediction server. Bioinformatics. 16:404-405.

45. Moss, P. 2022. The T cell immune response against SARS-CoV-2. Nature Immunology. 23:186-193.

46. Naveed, M., A.R. Yaseen, H. Khalid, U. Ali, A.A. Rabaan, M. Garout, M.A. Halwani, A. Al Mutair, S. Alhumaid, Z. Al Alawi, Y.N. Alhashem, N. Ahmed, and C.Y. Yean. 2022. Execution and Design of an Anti HPIV-1 Vaccine with Multiple Epitopes Triggering Innate and Adaptive Immune Responses: An Immunoinformatic Approach. In Vaccines. Vol. 10.

47. Nguyen, L.H., D.A. Drew, M.S. Graham, A.D. Joshi, C.-G. Guo, W. Ma, R.S. Mehta, E.T. Warner, D.R. Sikavi, and C.-H.J.T.L.P.H. Lo. 2020. Risk of COVID-19 among front-line health-care workers and the general community: a prospective cohort study. 5:e475-e483.

48. Oliveira, S.C., M.T.Q. de Magalhães, and E.J. Homan. 2020. Immunoinformatic Analysis of SARS-CoV-2 Nucleocapsid Protein and Identification of COVID-19 Vaccine Targets. 11.

49. Parvizpour, S., M.M. Pourseif, J. Razmara, M.A. Rafi, and Y. Omidi. 2020. Epitope-based vaccine design: a comprehensive overview of bioinformatics approaches. Drug Discovery Today. 25:1034-1042.

50. Petersen, E., F. Ntoumi, D.S. Hui, A. Abubakar, L.D. Kramer, C. Obiero, P.A. Tambyah, L. Blumberg, R. Yapi, and S.J.I.J.o.I.D. Al-Abri. 2022. Emergence of new SARS-CoV-2 Variant of Concern Omicron (B. 1.1. 529)-highlights Africa's research capabilities, but exposes major knowledge gaps, inequities of vaccine distribution, inadequacies in global COVID-19 response and control efforts. 114:268-272.

51. Polack, F.P., S.J. Thomas, N. Kitchin, J. Absalon, A. Gurtman, S. Lockhart, J.L. Perez, G. Pérez Marc, E.D. Moreira, C. Zerbini, R. Bailey, K.A. Swanson, S. Roychoudhury, K. Koury, P. Li, W.V. Kalina, D. Cooper, R.W. Frenck, Jr., L.L. Hammitt, Ö. Türeci, H. Nell, A. Schaefer, S. Ünal, D.B. Tresnan, S. Mather, P.R. Dormitzer, U. Şahin, K.U. Jansen, and W.C. Gruber. 2020. Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine. The New England journal of medicine. 383:2603-2615.

52. Rapin, N., O. Lund, M. Bernaschi, and F. Castiglione. 2010. Computational immunology meets bioinformatics: the use of prediction tools for molecular binding in the simulation of the immune system. PloS one. 5:e9862.

53. Reynisson, B., B. Alvarez, S. Paul, B. Peters, and M. Nielsen. 2020. NetMHCpan-4.1 and NetMHCIIpan-4.0: improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data. Nucleic acids research. 48:W449-w454.

54. Richner, J.M., S. Himansu, K.A. Dowd, S.L. Butler, V. Salazar, J.M. Fox, J.G. Julander, W.W. Tang, S. Shresta, T.C. Pierson, G. Ciaramella, and M.S. Diamond. 2017. Modified mRNA Vaccines Protect against Zika Virus Infection. Cell. 168:1114-1125.e1110.

55. Satarker, S., and M. Nampoothiri. 2020. Structural Proteins in Severe Acute Respiratory Syndrome Coronavirus-2. Archives of medical research. 51:482-491.

56. Sharma, N., L.D. Naorem, S. Jain, and G.P.S. Raghava. 2022. ToxinPred2: an improved method for predicting toxicity of proteins. Briefings in Bioinformatics. 23.

57. Treanor, J. 2004. Influenza vaccine--outmaneuvering antigenic shift and drift. The New England journal of medicine. 350:218-220.

58. Vankadari, N., J.A.J.E.m. Wilce, and infections. 2020. Emerging COVID-19 coronavirus: glycan shield and structure prediction of spike glycoprotein and its interaction with human CD26. 9:601-604.

59. Wang, C., P.W. Horby, F.G. Hayden, and G.F. Gao. 2020. A novel coronavirus outbreak of global health concern. Lancet (London, England). 395:470-473.

60. Wang, R., J. Chen, and G.-W. Wei. 2021. Mechanisms of SARS-CoV-2 Evolution Revealing Vaccine-Resistant Mutations in Europe and America. The Journal of Physical Chemistry Letters. 12:11850-11857.

61. Wiederstein, M., and M.J. Sippl. 2007. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic acids research. 35:W407-410.

62. Wlodawer, A. 2017. Stereochemistry and Validation of Macromolecular Structures. Methods in molecular biology (Clifton, N.J.). 1607:595-610.

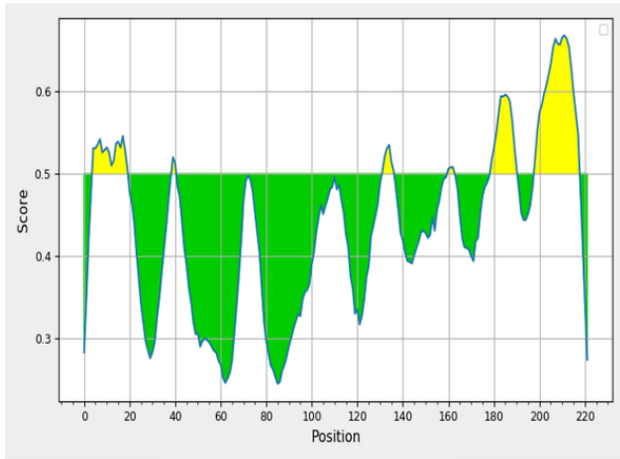63. Zhang, Y. 2008. I-TASSER server for protein 3D structure prediction. BMC Bioinformatics. 9:40.
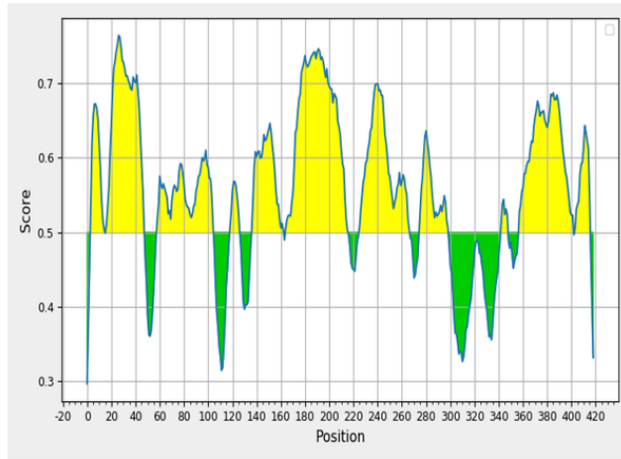
Figures
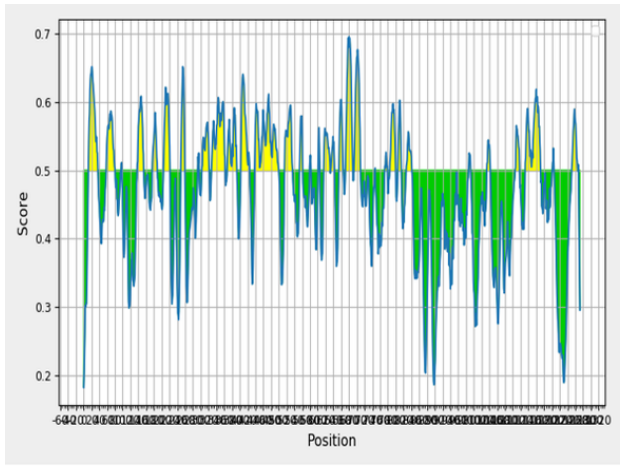
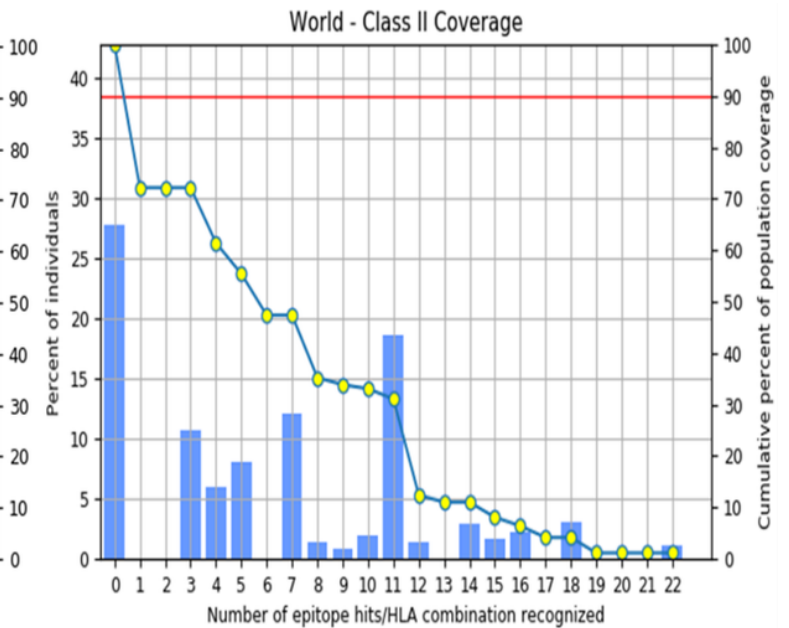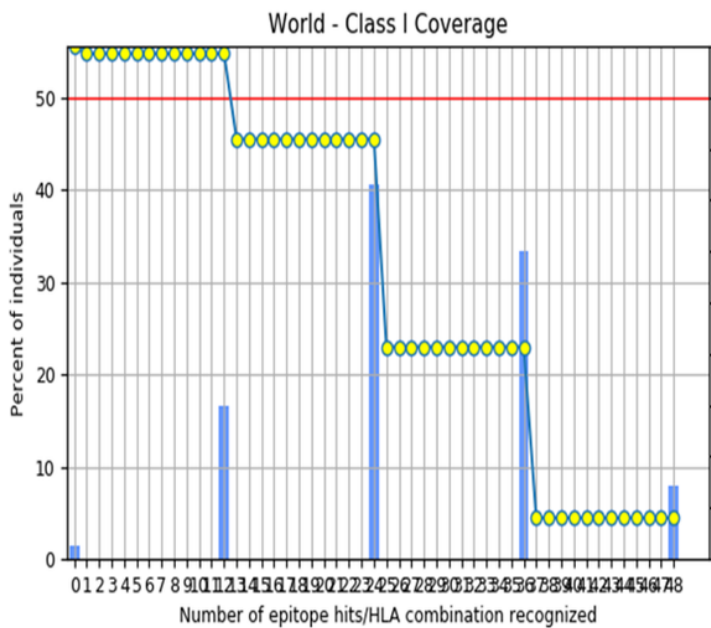**Spike Glycoprotein**
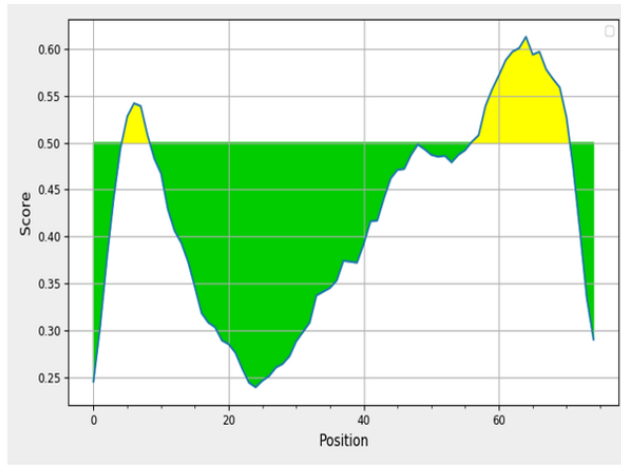
Envelope small membrane protein

Membrane protein

Nucleocapsid or Nucleoprotein

Figure 1

The graphs indicating the peptide regions (Yellow) that could become the part of epitope as their score is above the threshold (0.5), while the peptide regions below the threshold (Green) are not eligible to become a part of B-cell-restricted epitopes.

Figure 2

The graph on the left and right represent the cumulative world population coverage of selected MHC-I and MHC-II epitopes respectively in correspondence to the recognized number of HLA hits
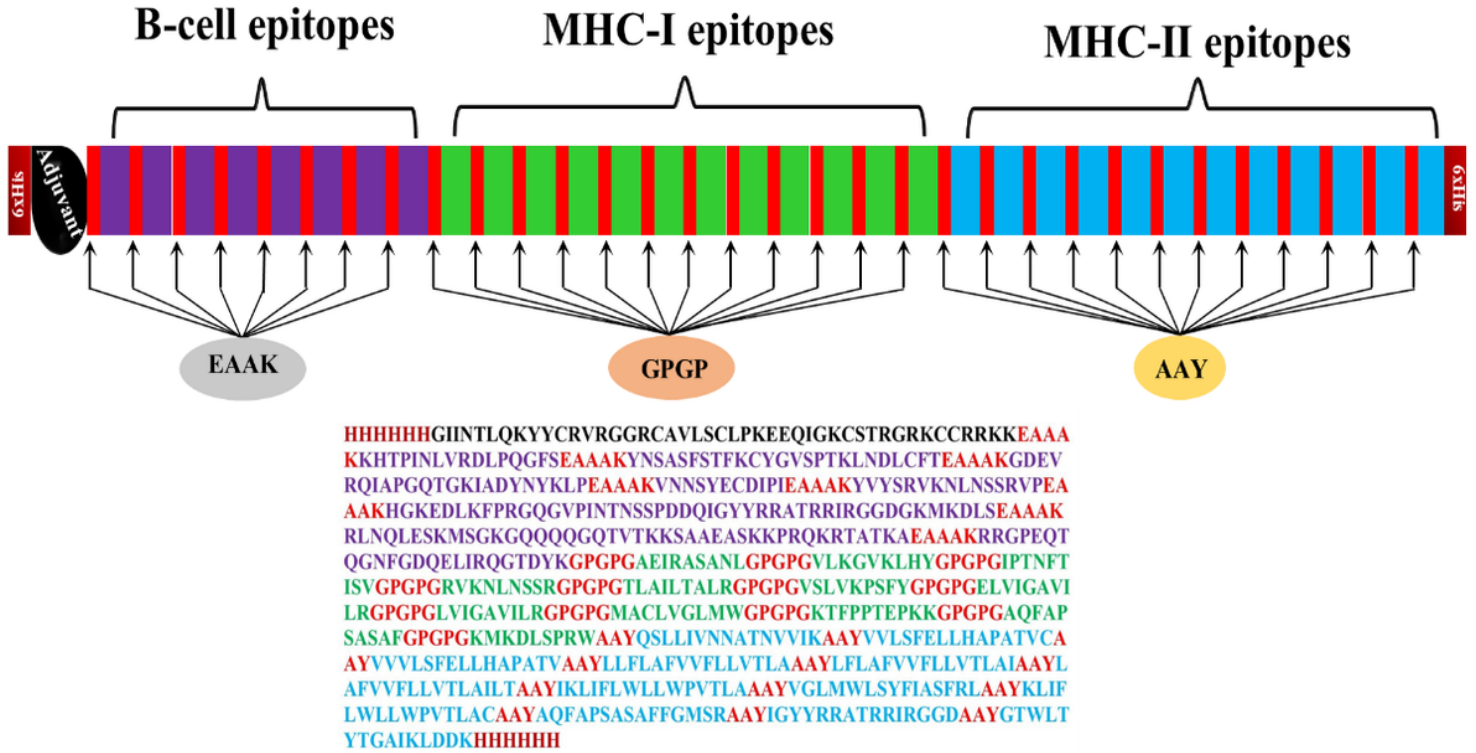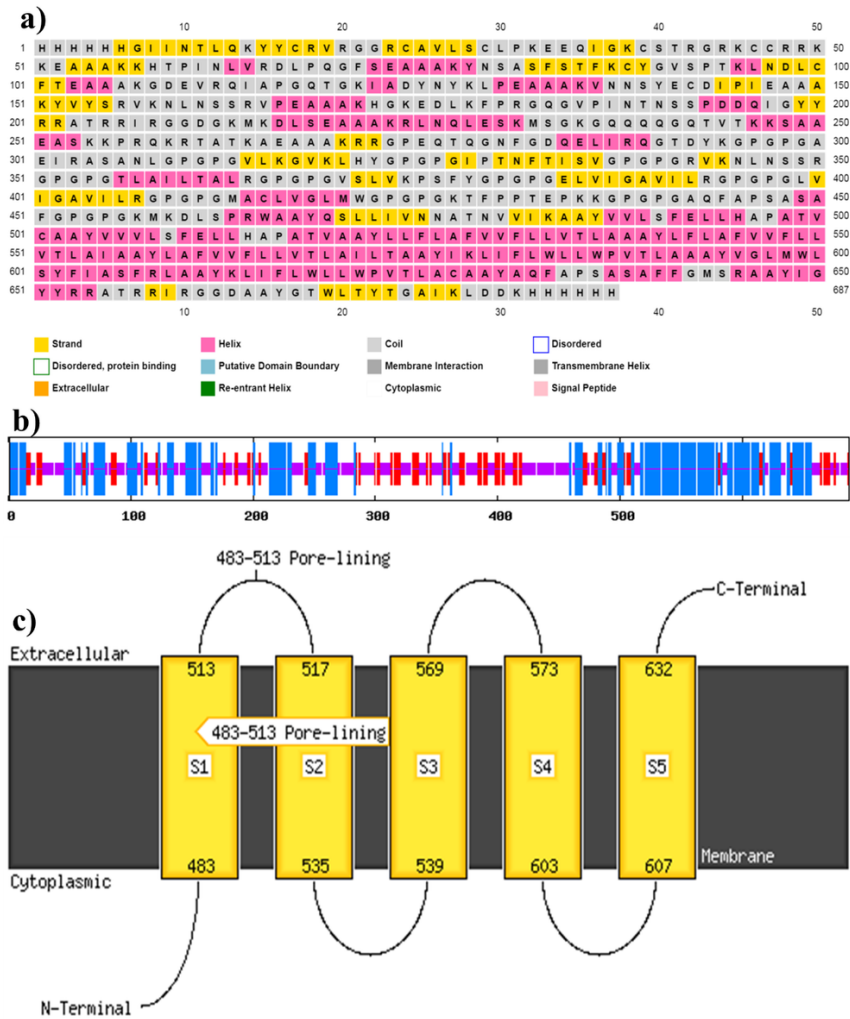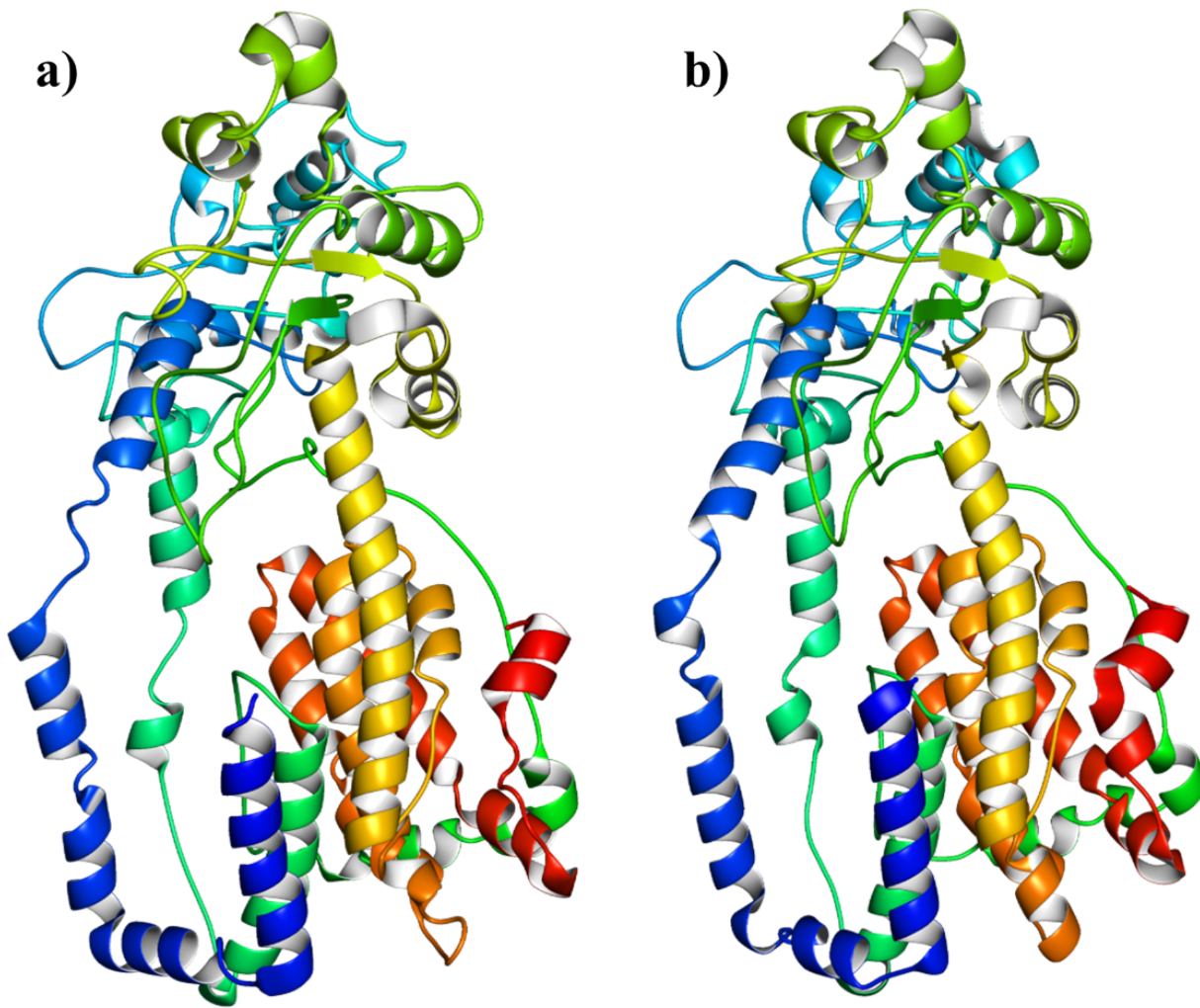

Figure 3

Graphical representation of vaccine assembly in sequential arrangements and the complete amino acid sequence of the designed chimeric vaccine

a)

	10	20	30	40	50
1	H H H H H H G I I N T L Q K Y Y C R V R G G R C A V L S C L P K E E Q I G K C S T R G R K C C R R K	50
51	K E A A A K K H T P I N L V R D L P Q G F S E A A A K Y N S A S F S T F K C Y G V S P T K L N D L C	100
101	F T E A A A K G D E V R Q I A P G Q T G K I A D Y N Y K L P E A A A K V N N S Y E C D I P I E A A A	150
151	K Y V Y S R V K N L N S S R V P E A A A K H G K E D L K F P R G Q G V P I N T N S S P D D Q I G Y Y	200
201	R R A T R R I R G G D G K M K D L S E A A A K R L N Q L E S K M S G K G Q Q Q Q G Q T V T K K S A A	250
251	E A S K K P R Q K R T A T K A E A A A K R R G P E Q T Q G N F G D Q E L I R Q G T D Y K G P G P G A	300
301	E I R A S A N L G P G P G V L K G V K L H Y G P G P G I P T N F T I S V G P G P G R V K N L N S S R	350
351	G P G P G T L A I L T A L R G P G P G V S L V K P S F Y G P G P G E L V I G A V I L R G P G P G L V	400
401	I G A V I L R G P G P G M A C L V G L M W G P G P G K T F P P T E P K K G P G P G A Q F A P S A S A	450
451	F G P G P G K M K D L S P R W A A Y Q S L L I V N N A T N V V I K A A Y V V L S F E L L H A P A T V	500
501	C A A Y V V V L S F E L L H A P A T V A A Y L L F L A F V V F L L V T L A A A Y L F L A F V V F L L	550
551	V T L A I A A Y L A F V V F L L V T L A I L T A A Y I K L I F L W L L W P V T L A A A Y V G L M W L	600
601	S Y F I A S F R L A A Y K L I F L W L L W P V T L A C A A Y A Q F A P S A S A F F G M S R A A Y I G	650
651	Y Y R R A T R R I R G G D A A Y G T W L T Y T G A I K L D D K H H H H H H	687
	10	20	30	40	50

Legend:
- Strand
- Helix
- Coil
- Disordered
- Disordered, protein binding
- Putative Domain Boundary
- Membrane Interaction
- Transmembrane Helix
- Extracellular
- Re-entrant Helix
- Cytoplasmic
- Signal Peptide

b)

c)

Figure 4

Secondary structure analysis of the construct: a) PsiPred results showing the distribution of various elements of secondary structure throughout the chain, b) Represent the SOPMA analysis for the calculation of residues in various elements (alpha helix, beta turns, coils, etc.), and c) Showed the MEMSAT-SVM schematic of the vaccine construct.

Figure 5

Three-dimensional model of the vaccine construct: a) Represent raw model with 78.3 residues in the most favored region, b) Showed the refined model with 91.1% residues in the most favored region
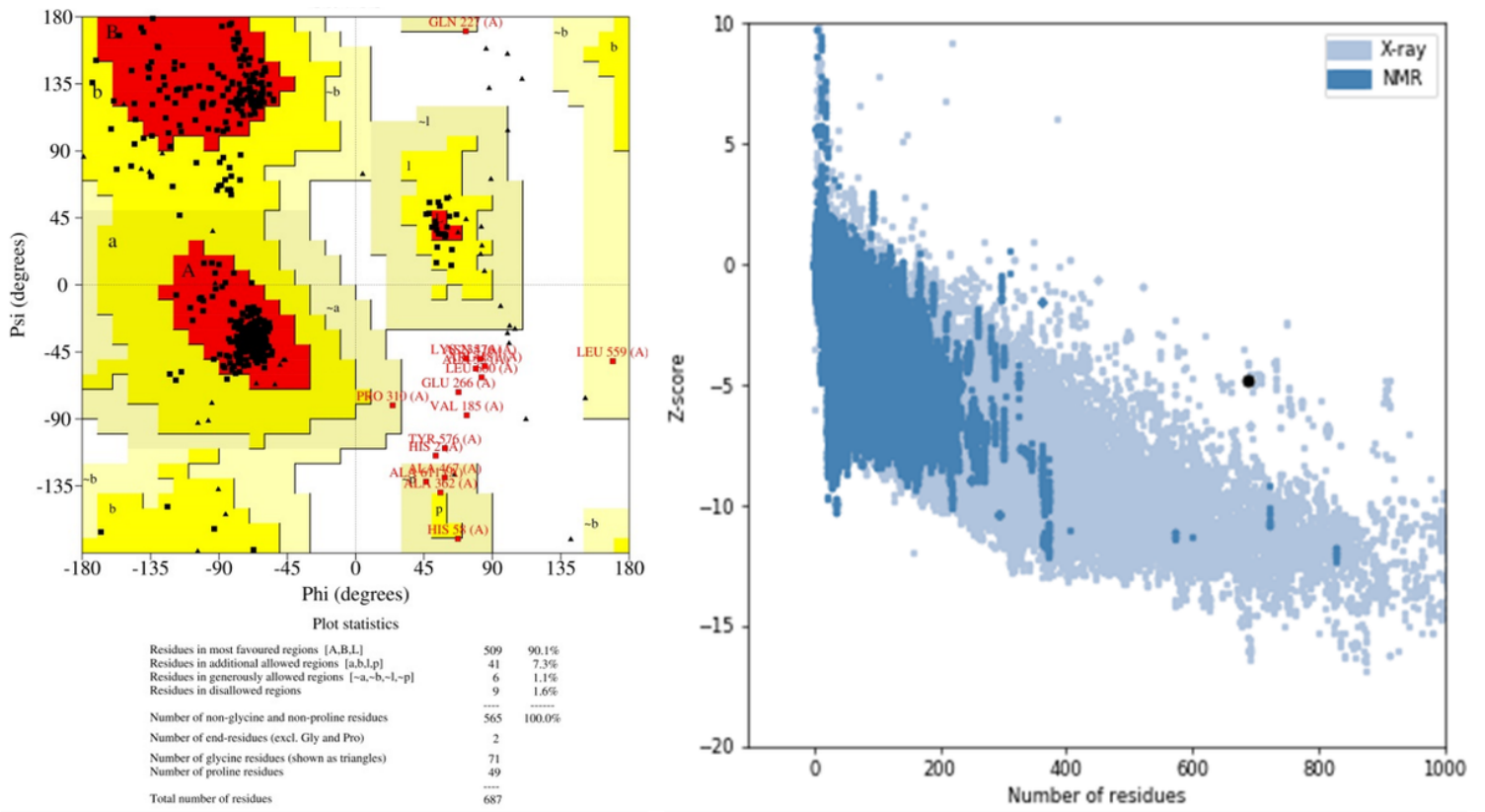
Figure 6

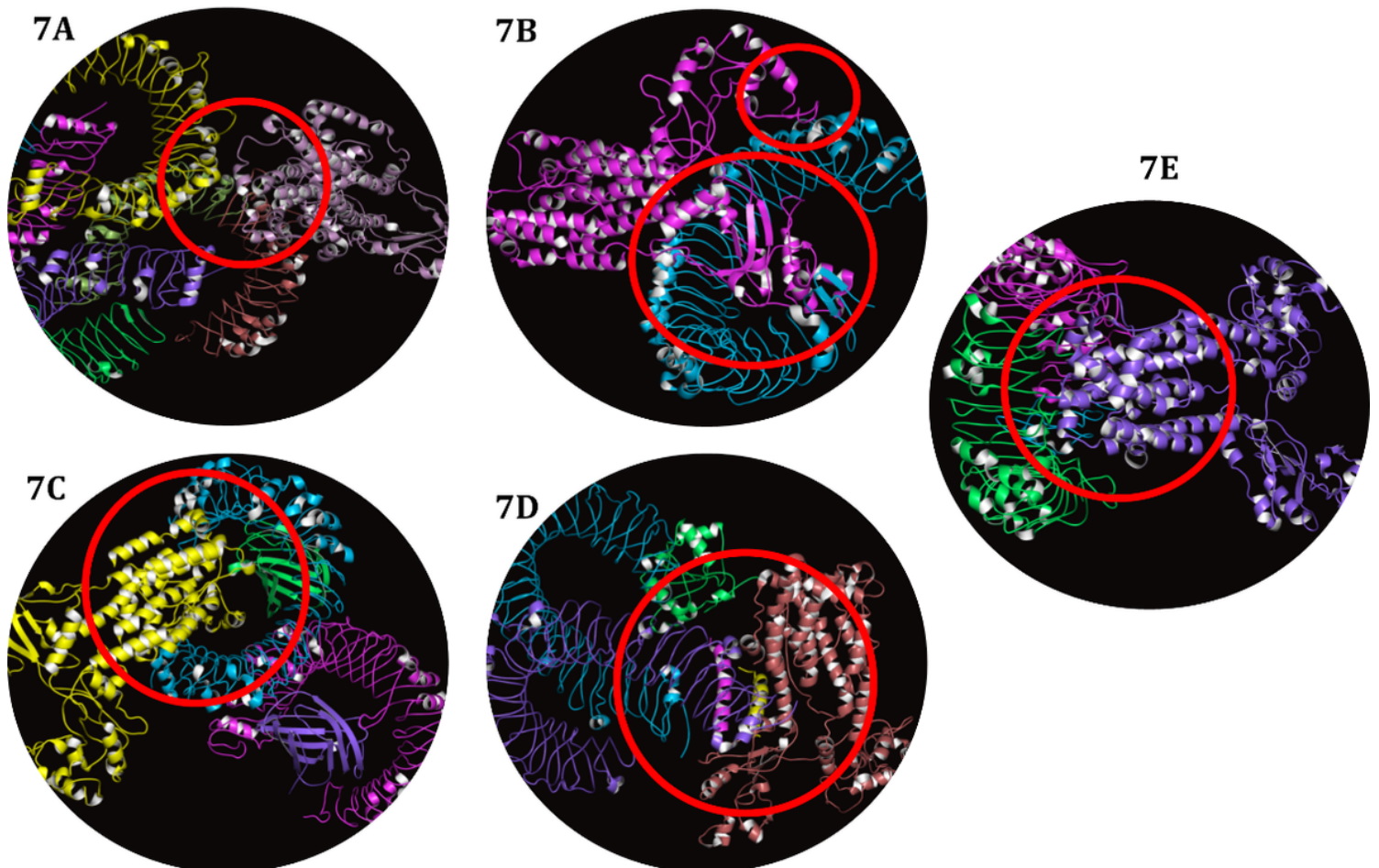The Ramachandran plot (left) and ProSA web analysis (right) for the validation of 3D model

Figure 7

The docked complexes between the vaccine construct and various TLRs (TLR2, TLR3, TLR4, TLR5, and TLR8)
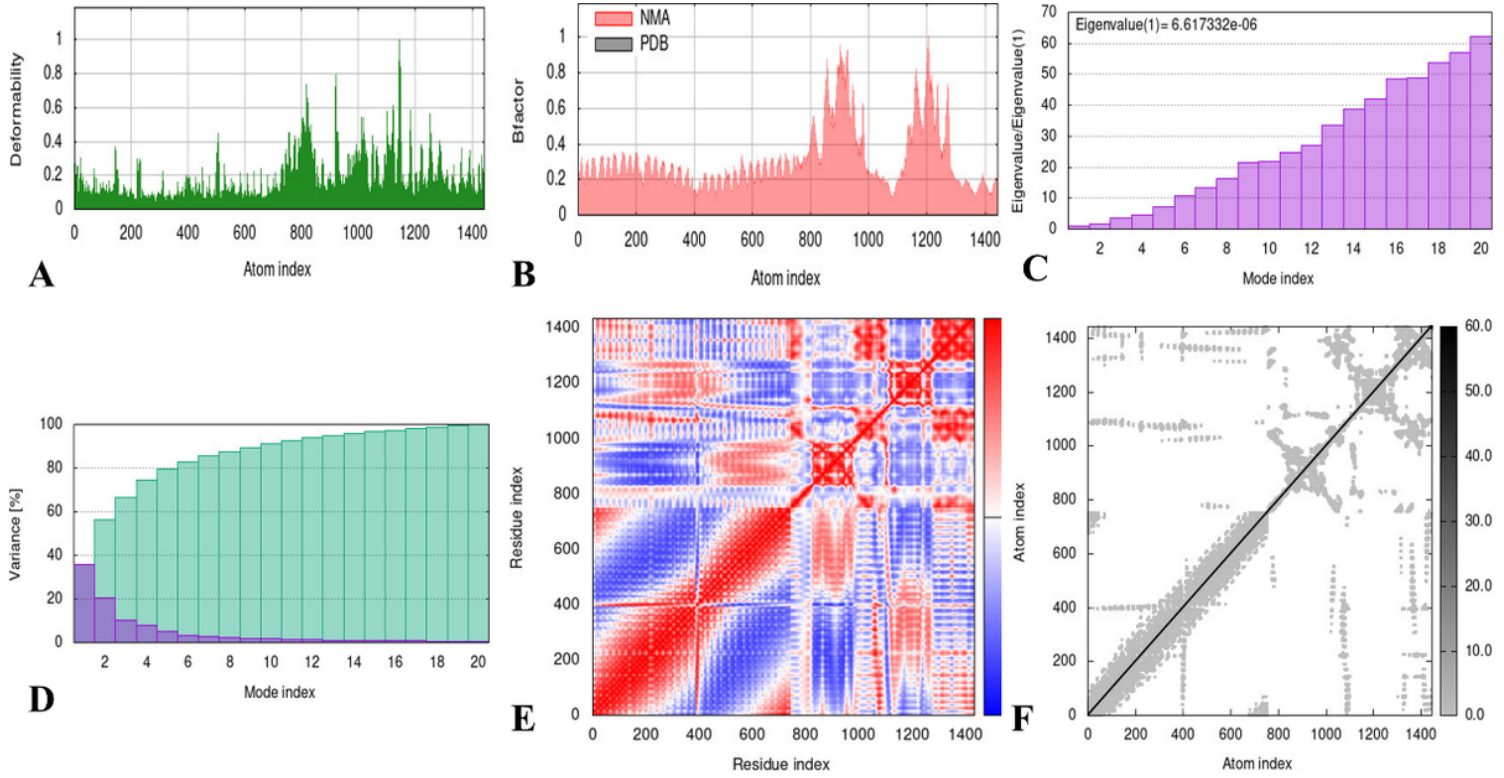

Figure 8

Graphical representation of various Normal mode Analyses (NMA) generated by iMODS for the docked complex (Vaccine + TLR8)
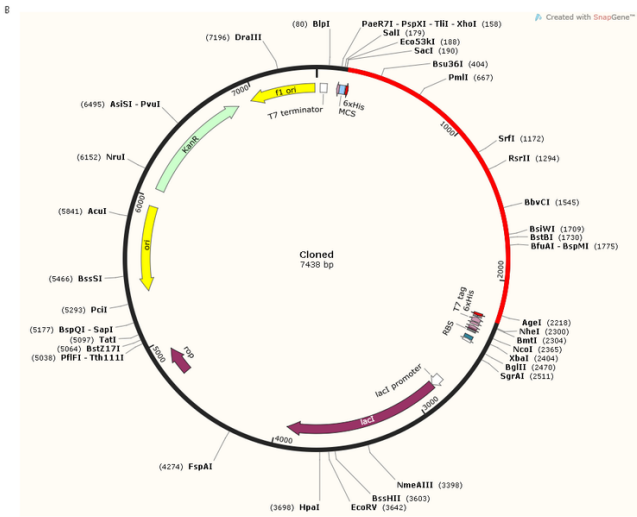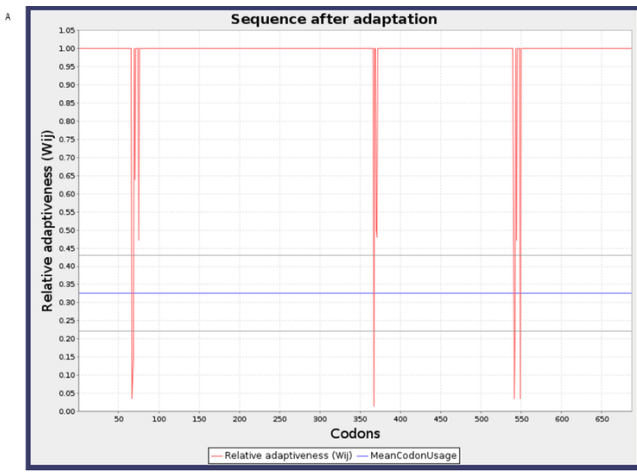
Figure 9

a: Codon adaption/optimization results generated by JCAT for the vaccine construct

b: The cloned vector generated by SnapGene containing the vaccine sequence inserted (represented in red color) in pET28a (+) plasmid
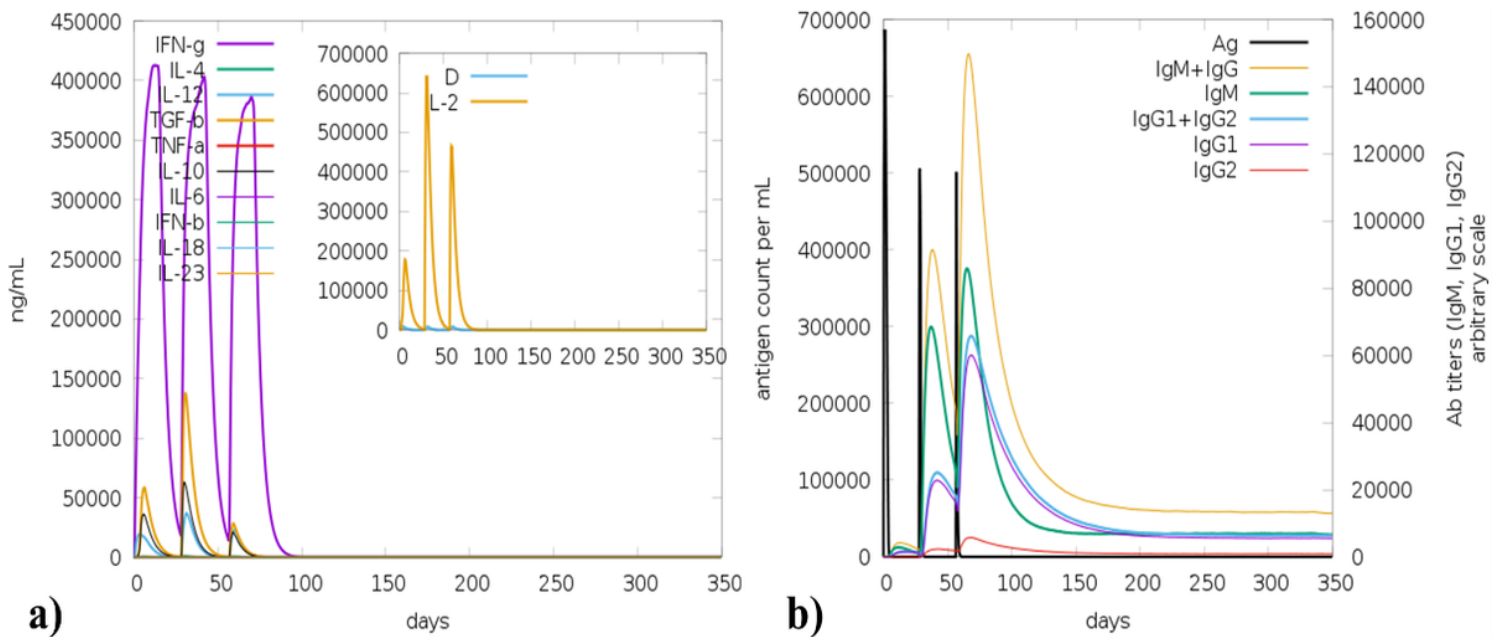

Figure 10

Immune simulation results: a) The sudden rise in the production of various cytokines can be observed while the sub-graph of the straight Danger (D) line ensures the safety of the vaccine, while b) represents the production of various antibodies.
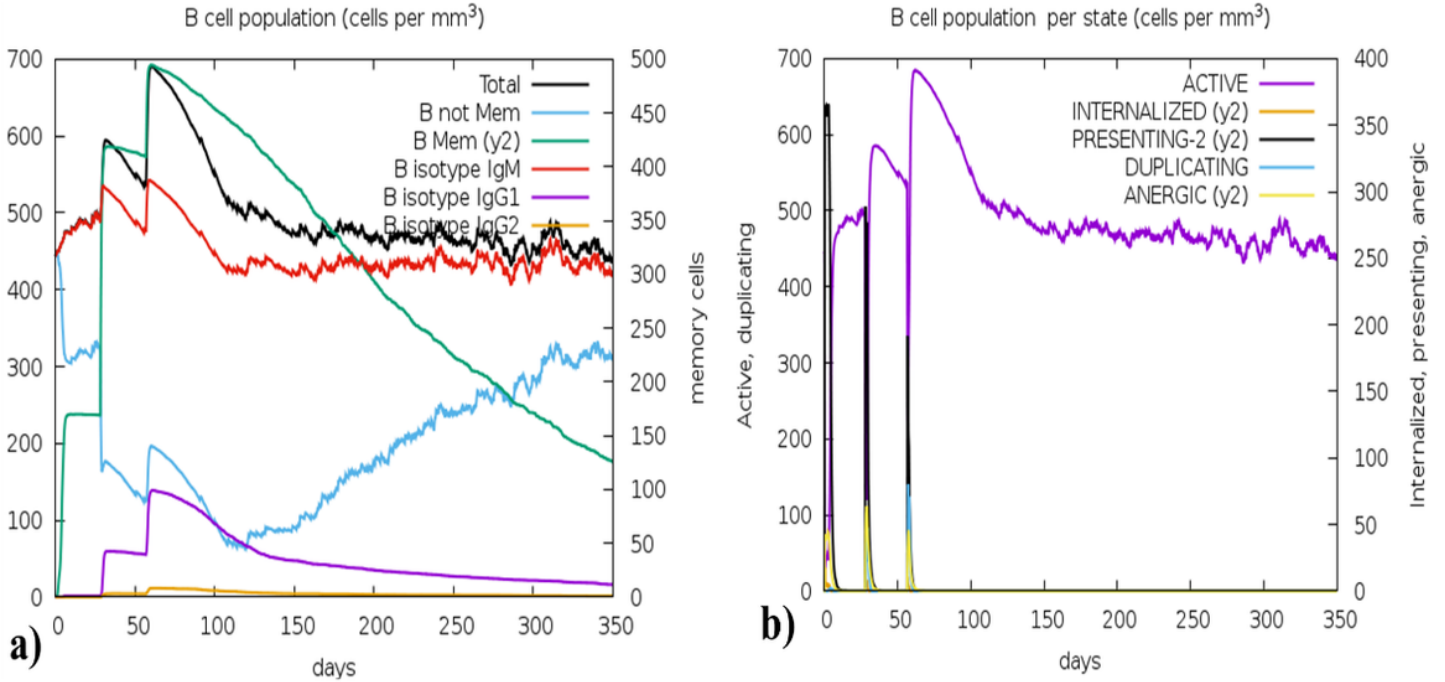

Figure 11

The immune simulation results, a) presents the production memory B-cells (mm-3) while b) represents the production B-cell population in various states (cells/mm3).
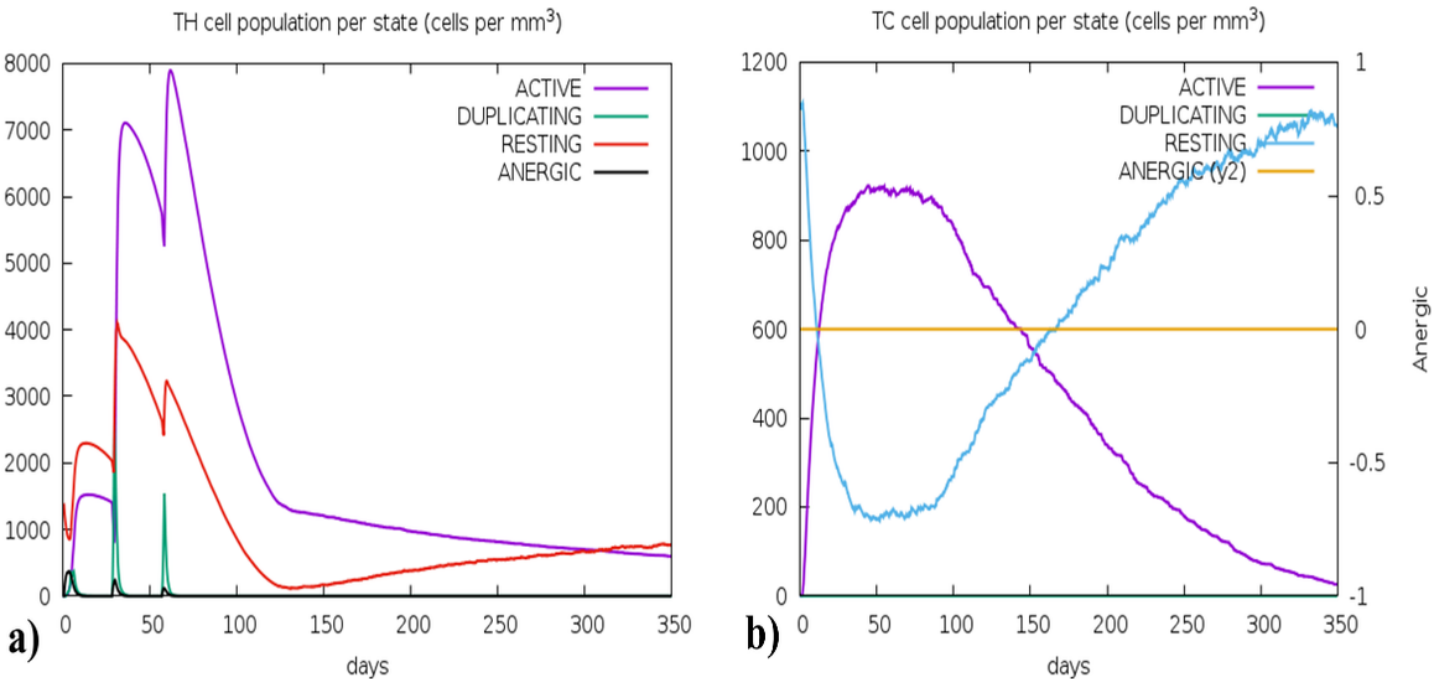

Figure 12

The immune simulation predictions by C-IMMSIMS, a) illustrate the production of helper T-cells per state (cells/mm3), and b) showed the Tc cell per state (cells/mm3).

Supplementary Files

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

- SuplementaryData.rar