RESEARCH ARTICLE

Proteome based analysis of circulating SARS-CoV-2 variants: approach to a universal vaccine candidate

Elijah Kolawole Oladipo[1,2] · Taiwo Ooreoluwa Ojo[2,3] · Seun Elijah Olufemi[2,3] · Boluwatife Ayobami Irewolede[2] ·
Daniel Adewole Adediran[2,3] · Asegunloluwa Grace Abiala[2,4] · Oluwaseun Samuel Hezekiah[2,4] ·
Akindele Felix Idowu[2,3] · Yinmi Gabriel Oladeji[2,8] · Mary Omotoyinbo Ikuomola[2,4] · Adenike Titilayo Olayinka[2,7] ·
Gideon Oluwamayowa Akanbi[2,5] · Usman Abiodun Idowu[2,5] · Odunola Abimbola Olubodun[2,4] ·
Folusho Daniel Odunlami[2,4] · James Akinwumi Ogunniran[2,7] · Omodamola Paulina Akinro[2,5] ·
Hadijat Motunrayo Adegoke[2,6] · Elizabeth Oluwatoyin Folakanmi[2,4] · Temitope Aishat Usman[2] ·
Elizabeth Folakemi Oladokun[2,5] · Glory Jesudara Oluwasanya[2] · Hezekiah Oluwajoba Awobiyi[2] ·
Jerry Ayobami Oluwasegun[2,4] · Samuel Adebowale Akintibubo[2,5] · Esther Moradeyo Jimah[2]

Abstract

The discovery of the first infectious variant in Wuhan, China, in December 2019, has posed concerns over global health due to the spread of COVID-19 and subsequent variants. While the majority of patients experience flu-like symptoms such as cold and fever, a small percentage, particularly those with compromised immune systems, progress from mild illness to fatality. COVID-19 is caused by a RNA virus known as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). Our approach involved utilizing immunoinformatic to identify vaccine candidates with multiple epitopes and ligand-binding regions in reported SARS-CoV-2 variants. Through analysis of the spike glycoprotein, we identified dominant epitopes for T-cells and B-cells, resulting in a vaccine construct containing two helper T-cell epitopes, six cytotoxic T-cell epitopes, and four linear B-cell epitopes. Prior to conjugation with adjuvants and linkers, all epitopes were evaluated for antigenicity, toxicity, and allergenicity. Additionally, we assessed the vaccine Toll-Like Receptors complex (2, 3, and 4). The vaccine construct demonstrated antigenicity, non-toxicity, and non-allergenicity, thereby enabling the host to generate antibodies with favorable physicochemical characteristics. Furthermore, the 3D structure of the B-cell construct exhibited a ProSA-web z-score plot with a value of -1.71, indicating the reliability of the designed structure. The Ramachandran plot analysis revealed that 99.6% of the amino acid residues in the vaccine subunit were located in the high favored observation region, further establishing its strong candidacy as a vaccination option.

Keywords Covid-19 · SARS-Cov-2 · Variant · Vaccine · Immunoinformatic · Multi-epitope

✉ Elijah Kolawole Oladipo
  koladipo2k3@yahoo.co.uk

1 Department of Microbiology, Laboratory of Molecular Biology, Immunology and Informatics, Adeleke University, Ede, Osun State, Nigeria

2 Genomics Unit, Helix Biogen Institute, Ogbomoso, Oyo State, Nigeria

3 Department of Biochemistry, Ladoke Akintola University of Technology, Ogbomoso, Oyo State, Nigeria

4 Department of Physiology, Ladoke Akintola University of Technology, Ogbomoso, Oyo State, Nigeria

5 Department of Pure and Applied Biology, Microbiology Unit, Ladoke Akintola University of Technology, Ogbomoso, Oyo State, Nigeria

6 Computational Biophysical Chemistry Laboratory, Department of Pure and Applied Chemistry, Ladoke Akintola University of Technology, Ogbomoso, Oyo State, Nigeria

7 Department of Medical Microbiology and Parasitology, Ladoke Akintola University of Technology, Ogbomoso, Oyo State, Nigeria

8 Department of Microbiology, Obafemi Awolowo University, Ile Ife, Osun State, Nigeria

Introduction

Since the onset of the COVID-19 pandemic in 2020, the global community has been grappling with the enduring impact of this highly burdensome and formidable disease. Renowned health organizations such as the Centre for Disease Control (CDC) and the World Health Organization (WHO) have been tirelessly engaged in efforts to mitigate its devastating consequences. SARS-CoV-2, initially identified as 2019-nCov, emerged as a virus of immense global significance, originating in Wuhan, China. This viral pathogen exerts a particularly severe toll on vulnerable populations, including the elderly and children, with documented symptomatic cases exhibiting distressing symptoms such as fever, dry cough, dyspnea, and distinct bilateral ground-glass opacities on chest CT scans (Chandra 2012; Nandakumar 2020; Sun et al. 2020).

However, during the initial emergence of the first strain of SARS-CoV-2, there existed no medically validated immunoprophylactic agent to combat the infection (Chandra 2012). Notably, in China, the country of origin for the causative agent, various interventions were employed, including the administration of antimalarial drugs like Chloroquine and Hydroxychloroquine, antivirals such as Lopinavir and Ritonavir, and convalescent patient serum but to no avail. Therefore, as part of the global concerted effort to curtail the widespread transmission of SARS-CoV-2, endeavors have been undertaken to develop vaccines capable of eliciting adaptive immunity in humans, consequently facilitating herd immunity within the population (Resolution & Others 2015). These interventions aimed to eliminate SARS-CoV-2 from infected individuals and confer a certain degree of immunity during the nascent stage of the pandemic (Nandakumar 2020).

Sadly, even the vaccines manufactured in the contemporary times to confer significant acquired adaptive and active immunity against the virus have been faced with limitations because of the several strains and variants of the RNA virus that have evolved as a result of mutation. Due to the nucleic acid of the virus being RNA, the mutation rate of the virus is extremely high, consequently, new strains and variants of the virus are constantly emerging (Sanjua 2016). With the highly pronounced diversity of the SARS-CoV-2, however, about 6.7 billion people have been vaccinated globally with any of the available vaccines which is medically recommended based on age differences and the likelihood of allergenicity as of October 2021 (Myers 2021). Therefore, there is as such, no universal Vaccine available for use as an agent that is able to work against all of the strains of SARS-CoV-2 and their individual variants. For example, the interim analysis of phase 3 ENSEMBLE trial of the Janseen's COVID-19 Vaccine candidate showed a lower efficiency in countries other than the United States, especially those countries where there is a higher rate of SARS-CoV-2 virus variants, though it was 72% effective in the United States (pharmacypracticenews.com). The need for an immunoprophylactic agent that is universal in action is inevitably needed to confer the required acquired immunity to all people of all ages in all locations in the world.

The design of a universal vaccine that is expected to provide immunity against every variation of SARS-CoV 2 and each strain is presented in this research and also spike glycoprotein (S), the site of the primary mutations that produced these diverse SARS-CoV-2 virus variants was used as the basis for design using a variety of immunoinformatic tools.

Methodology

Data assembly and retrieval of sequence

In order to reflect the circulating variants on the various continents of the world, representative SARS-CoV-2 whole-genome sequences were acquired from the Global Initiative on Sharing all Influenza Data (GISAID) databanks. We considered all of the World Health Organization's Variants of Concern (VOC) and Variants of Interest (VOI) (Oladipo et al. 2021).

An average of twenty (20) sequences were retrieved per variant across atleast five (5) out of six (6) continents. Sequences of variants of concern and interest which include B.1.214.2, B.1.1.7, B.1.525, B.1.526, B.1.351, B.1.1318, B.1.1328, B.1.1.519, B.1.621, R.1, B.1.4662, C.37, P.1, C.1.2, B.1.617.1, B.1.427, C.36.3, B.1.427, B.1.620, B.1.617.2, B.1.619, and B.1.529 across various countries of Oceania, Africa, north America, South America, Europe and Asia, were used for the analysis and vaccine design (Oladipo et al. 2020).

Annotation of protein sequences

Protein transeq was used to convert the SARS-CoV-2 nucleotide sequences obtained from GISAID from 5 continents to protein sequences. Using MEGA-X software (CLUSTA-W tool) to map out the annotated region, each complete protein sequence was annotated to determine the site of the previously established antigenic spike glycoprotein of SARS-CoV-2 (the first SARS-CoV-2 sequence from WUHAN) (Kumar et al. 2018).

CTL prediction

Using the NetCTL 1.2 server, CTL epitope predictions were generated (Madlala et al. 2021). CTL prediction

Fig. 1 Construction of the multi-epitope vaccine

was performed using SARS-CoV-2 sequences that were retrieved, annotated with reference spike glycoprotein sequence and also assessed through the antigenicity prediction test via the Vaxijen server, the CTL epitopes prediction threshold value for epitope identification was 0.75; the weights for C-terminal cleavage were 0.15; and the TAP transport efficiency was 0.05. The MHC class I binding prediction is expanded in the server version 1.2 to include 12 MHC supertypes, including the supertypes A26 and B39. The three main factors—MHC-1 binding peptides, proteasomal C-terminal cleavages, and Transporter Associated with Antigen Processing (TAP) transport efficiency—were used to underpin the prediction (Rapin and Wiernsperger. 2010).

Prediction of helper T cell (HTL) epitopes

Helper T cells (HTL) of the SARS-CoV-2 variants were predicted using Immune Epitope Database (IEDB), (Vita et al. 2015). Seven Human Alleles were selected using the IEDB recommended human allele which include HLA-DRB5*01:01, HLA-DRB1*03:01, HLA-DRB1*15:01, HLA-DRB1*07:01, HLA-DRB3*01:01, HLA-DRB3*02:02, HLA-DRB4*01:01 are selected for Major Histocompatibility Complex II (MHC II). The predicted epitopes were selected in the order of the smallest percentile which has a relationship with the MHC II binding (Zhang et al. 2008). Seven epitopes were selected for human allele and forty-nine for each variant.

Prediction of B-cell epitope

The SARS COV-2 glycoprotein's linear B-cell epitopes were predicted, and epitopes that exceeded 0.5 thresholds were chosen and subjected to additional analysis (Oladipo et al. 2020). For the prediction of a linear B-cell epitope coupled with a conformational B-cell epitope, we use BCPred and ABCpred. The kernel method's allotted support vector

machine method for amino acid pair antigenicity (AAP). While ABCPred forecasts a variety of pathogenic groups and was used to confirm BCPred's predictions. But it is claimed that residues with a wide range of values have better solvent accessibility (Ali et al. 2022).

Multi-epitope vaccine construction

Several B-Cells epitopes were discarded and only four (4) were used for the candidate vaccine construction. This was due to their poor performance in antigenicity tests, allergenicity prediction, and toxicity testing. Two HTL-epitopes were finally used for the vaccine construct. Initially, 116 HTL-epitopes were predicted but others were rejected based on the prediction parameters such as IL-4, IL-10, and IFN. CTL region is constructed using six (6) subunits of forty-four (44) which have improved performance in the CTL epitopes antigenicity prediction. The final vaccine candidate contains an adjuvant, human beta defensin-1 (accession no. P60022) (Ryan and Diamond 2017), to augment immune response (Yang et al. 2021) and the sequence was retrieved from the Uniprot, linked to the vaccine construct through an EAAAK linker, connected to the amino (N) terminus of the multi-subunit sequence (Rehman et al. 2021) to the first B-cell epitope. GPGPG linker was used to fuse subsequent B-cell epitopes. In between the last B-cell epitope and the first HTL epitope, GPGPG linker was used to linking the epitopes and was used to merge the subsequent HTL epitope (Shey et al. 2019). The CTL epitopes were linked using AAY linker (Oladipo et al. 2021). The positioning of the epitopes, adjuvant, and linkers that make up the vaccine construct is depicted in Fig. 1 below.

Fig. 2 The flow chart of the methodology

Toxicity, allergenicity, antigenicity, and solubility prediction

The goal of vaccination is to elicit an immunological response in the recipient. As a result, the potential vaccine should be stable, highly antigenic, non-allergic, non-toxic, and have suitable solubility (Pyasi et al. 2021). Physicochemical characteristics and solubility qualities of vaccine candidates aid in determining the vaccine's efficiency and effectiveness (Id et al. 2021a, b). The toxicity of the final multi-epitope vaccine sequence was evaluated by the ToxinPred (Chen et al. 2020a, b). We also evaluate the allergenicity of the final vaccine construct using Aller-TOP v.2.0 (Madlala et al. 2021).The antigenicity of the vaccine candidate was predicted using Vaxijen 2.0 online server using a threshold of 0.5 (Ebrahimi et al. 2019) and the AntigenPro (Yang et al. 2021).The antigenicity of the

constructs was predicted using an antigenic probability of ≥ 0.8 (Onile et al. 2020). The Vaxijen server was utilized to investigate the antigenicity probability of the protein utilizing a non-alignment-based method that focuses on the physiochemical characteristics of the prototype vaccine (Oladipo et al. 2021). The solubility of the final vaccine was finally evaluated by SolPro (Yang et al. 2021) and Protein-sol server (Rawal et al. 2021). Both of these servers rely on machine learning to provide significantly more accurate predictions. All parameters were retained at their default values during prediction (Sarkar et al. 2020).

Physicochemical property

The physicochemical qualities are crucial in determining how the vaccine interacts with its physiological

environment (Yang et al. 2021). ExPASy ProtParam Tool predicts physicochemical parameters such as hydropathicity, charge, half-life, instability index, pI (theoretical isoelectric point value), and molecular weight (Doytchinova and Flower 2008).
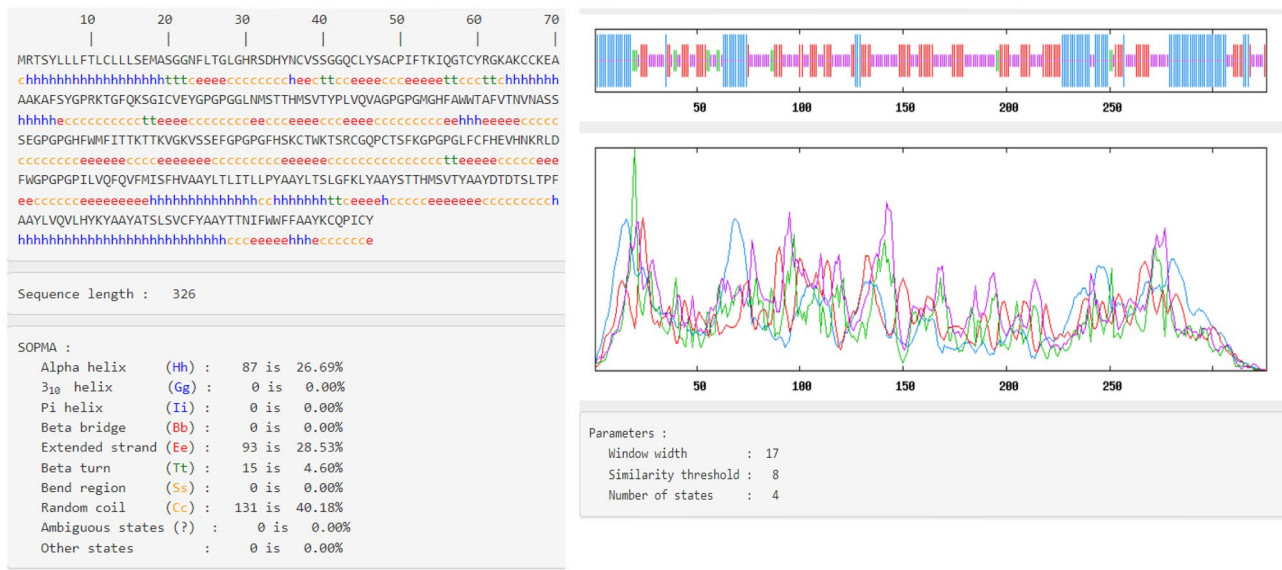
Secondary structure prediction

Using primary amino acid sequence input, the web-based server PSIPRED was used to accurately predict the protein secondary structure (Shey et al. 2019). PSI-BLAST (Position-Specific Iterated—BLAST) output is analyzed by PSIPRED using two feed-forward neural networks. It can reliably recognize secondary structure, as seen by its average Q3 score of 81.6% (Yang et al. 2021). The secondary structure of the vaccine construct was predicted using the Self-Optimized Prediction Method (SOPMA). One of the parameters predicted by SOPMA are helices, sheets, twists, and coils (Oladipo et al. 2020) set at default (Singh et al. 2020). The method makes use of customized BLAST to identify and select sequences that clearly resemble the intended vaccine (Oladipo et al. 2021).
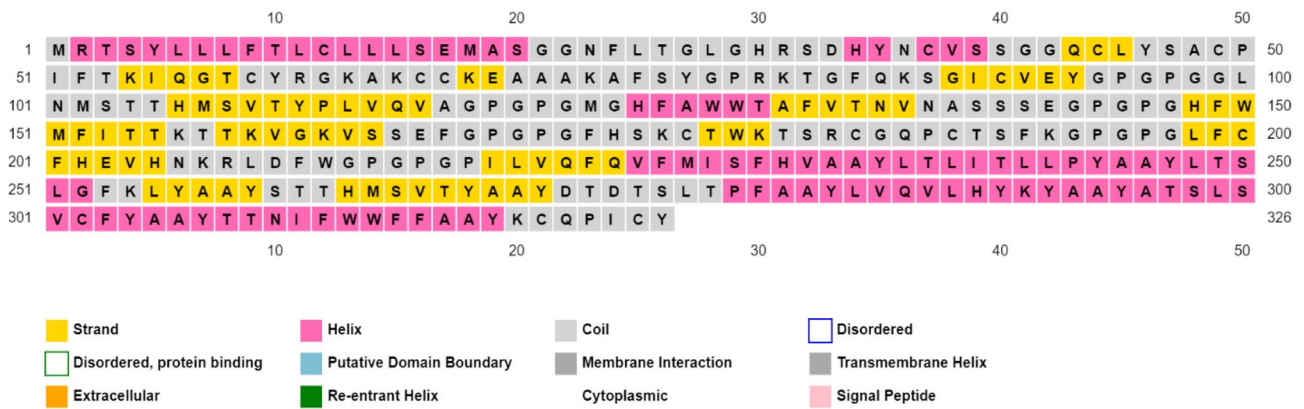
3D structure prediction

The 3D structure of target proteins was modeled using a modeling tool, 3Dpro of Scratch Protein predictor. This tool predicts the three-dimensional structure of the protein based on multiple-threading alignments and iterative prototype fragment assembly simulations. (Rehman et al. 2021).

## Structure refinement and validation

Computationally created 3D structures may not represent the real or native structures of proteins. As a result, 3D structure refinement is used to improve the resolution of computationally predicted models so that they mimic native protein structures as nearly as possible (Sarkar et al. 2020). The refinement of the 3D structure of the chimeric vaccine construct was performed by the GalaxyRefine tool of Galaxy-Web and desirable properties were obtained. After refinement, the GalaxyRefine tool output a total of five vaccination chimera models (Naveed et al. 2021). This server, which can improve both local and global protein structure quality, rebuilds and repacks the protein side-chain using a community-wide CASP10-based refinement approach that has been successfully tested (Onile et al. 2020). Subsequently, the best model (model 1) from Galaxyrefine was validated using ProSA-web (Adhikari et al. 2018) which provides a z-score. A z-score that falls inside the range of all the experimentally determined protein chains in the PDB database indicates that a query protein is of higher quality (Sarkar et al. 2020). Potential inaccuracies in the anticipated tertiary structure can be detected using the ProSA-web structure. The server displays the overall quality score as well as any problematic sections of the protein structure (Onile et al. 2020). Finally, the overall structural quality was validated by a Ramachandran plot analysis (Mahmud et al. 2021). A Ramachandran plot is a graphic representation of an amino acid's main-chain conformational tendencies (Anderson et al. 2005).
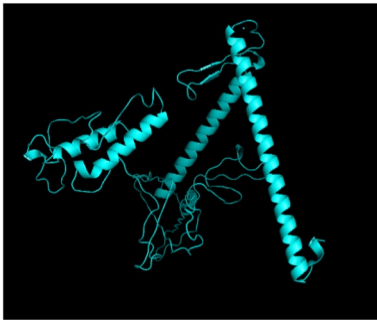

Fig. 3 Secondary structure of the vaccine construct as predicted by SOPMA

Fig. 4 Secondary structure of the protein construct as predicted by PSIPRED
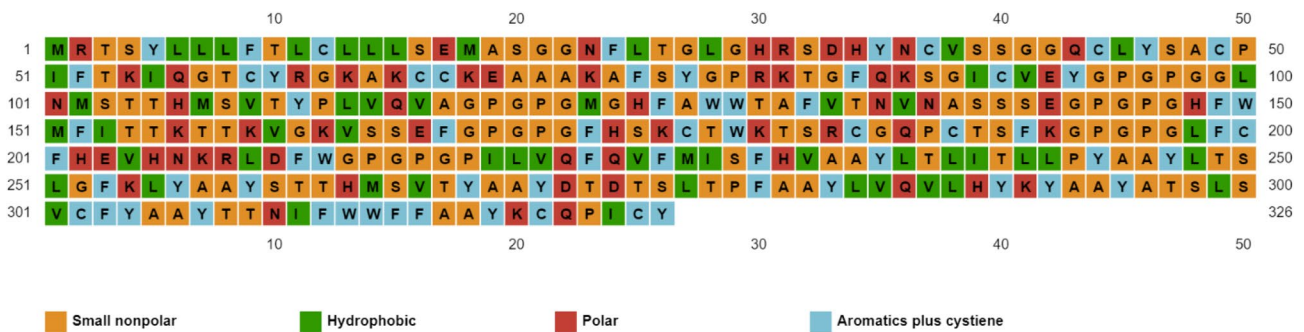
Fig. 5 Other parameters of the protein construct as predicted by PSIPRED

**Fig. 6** 3D-Structure of the vaccine construct viewed with pymol

Conformational B cell epitope prediction

Continuous and linear epitopes are the two basic types of B-cell epitopes discontinuous or conformational B-cell epitope on B-cell (Adhikari et al. 2018). ElliPro at IEDB was utilized for identifying the conformational B-cell epitopes. The vaccine protein's PDB structure was utilized as an input for predicting conformational epitopes in the vaccine construct's tertiary structure (Id et al. 2021a, b). ElliPro is a web-based tool that accepts two types of data as input: protein sequence and structure. With threshold values for BLAST e-value and the number of structural templates from PDB, the first example uses a protein SwissProt/UniProt ID or a sequence in either FASTA format or single letter codes to generate a 3D structure from the provided sequence (Ponomarenko et al. 2008). On the vaccine's three-dimensional structure, ElliPro discovered eleven (11) epitopes.

Codon optimization of designed vaccine peptide for expression analysis

To enable maximal protein production, the Java codon adaptation tool (JCat) was used to optimize codons (Naveed et al. 2021). This procedure was largely carried out using the Java Codon Adaptation Tool (JCat) server, an adaptor bioinformatics tool (Shey et al. 2019). The proportion of the C-G ratio was calculated using codon optimization as a strategy. A nucleotide with a higher C-G ratio has a better likelihood of being constructed. Codon optimization increases protein and RNA levels significantly, implying that codon use is an essential determinant of gene expression (Zhou et al. 2016).
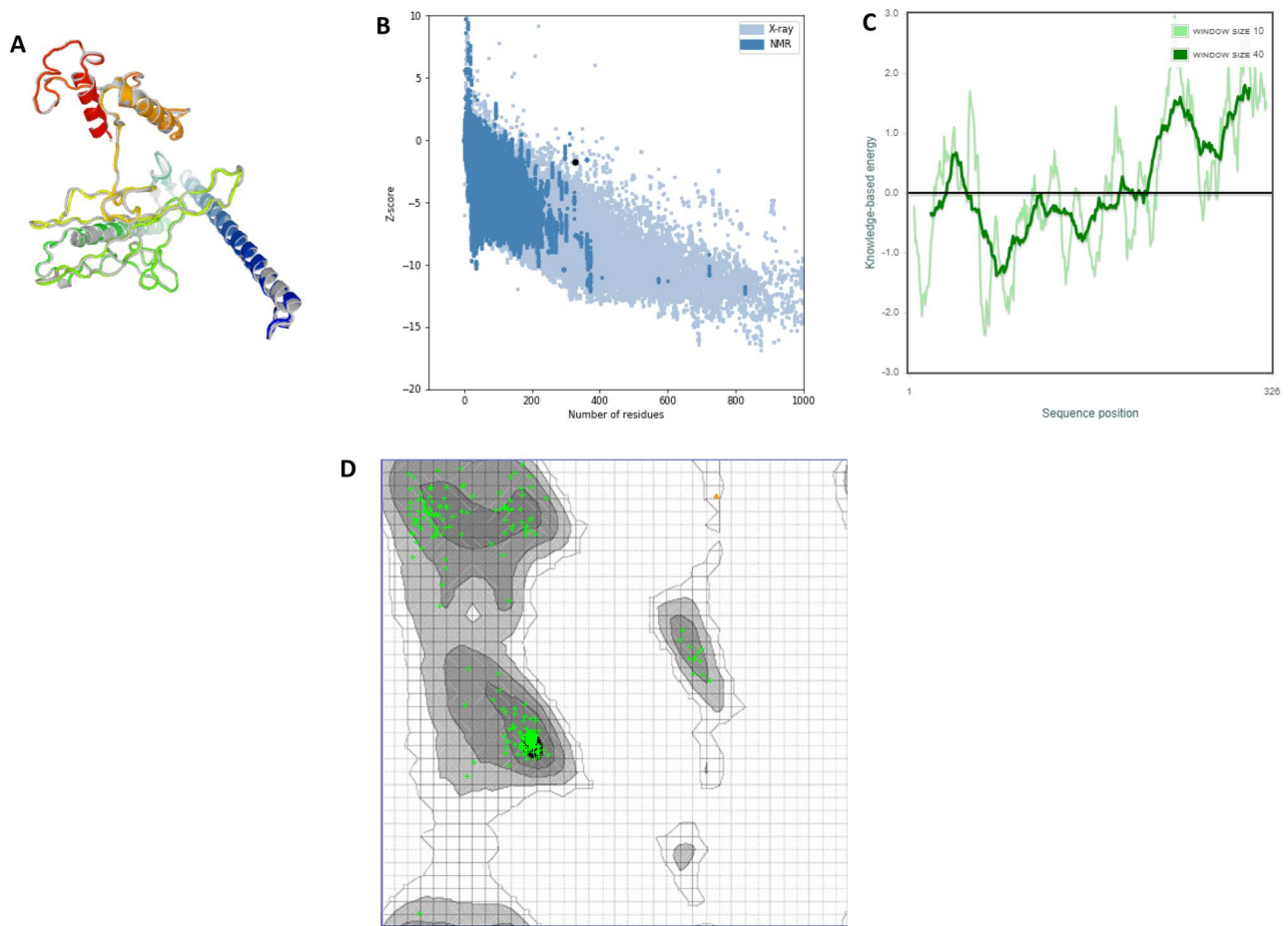
Fig. 7 3D-Structure refinement and validation structure. **A** Galaxy refined structure, **B** Overall model quality **C** Local model quality, **D** Ramachandran plot graph

Table 1 Structure information of the five (5) proposed model

| Model	GDT-HA	RMSD	MolProbity	Clash score	Poor rotamers	Rama favored
Initial	1.0000	0.000	3.820	119.3	8.8	89.5
Model 1	0.9041	0.537	1.981	15.0	0.8	95.7
Model 2	0.9095	0.522	1.945	15.6	0.4	96.3
Model 3	0.8926	0.547	1.940	14.4	0.4	96.0
Model 4	0.9018	0.527	1.967	15.4	0.4	96.0
Model 5	0.9011	0.536	1.964	14.2	0.8	95.7

JCat's result is presented graphically as well as Codon Adaptation Index (CAI) values for the pasted and newly adapted sequences (No et al. 2005). A sequence's C-G composition should be between 30 and 70 percent. Outside of this range, C-G content has negative implications for translational and transcriptional efficiency (Shey et al. 2019).

Molecular docking of vaccine construct with cleaned TLR 2, 3, and 4 receptors

TLRs play a critical role in the innate identification of invading microbes and their subsequent clearance. Toll-like receptor 2, 3 and 4 (TLR 2,3 and 4) was chosen as the receptor and retrieved from the RCSB PDB database (PDB ID: 5d3i, 7c76, 2z64 respectively), while the vaccine model was used as a ligand. Due to the fact that TLRs retrieved from the PDB database contain other ligands and water molecules

Table 2 Conformational B-Cell epitope prediction of the Vaccine construct

No	Start	End	Peptide	Number of residues	Score
1	301	326	VCFYAAYTTNIFW-WFFAAYKCQPICY	26	0.872
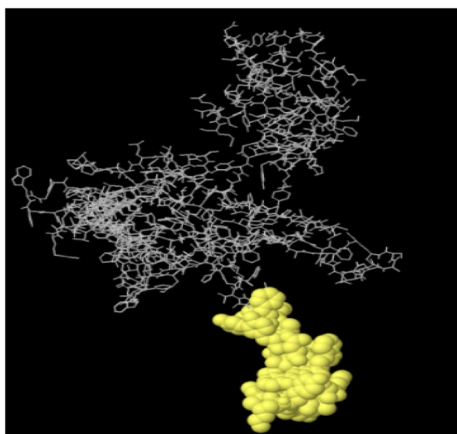
Fig. 8 Structure showing the position of the preferred conformational B-Cell epitope

that could interfere with the vaccine construct's efficient binding to the receptor's pocket, Discovery studio software was used to clean the protein and remove any extra ligands before docking. Using the ClusPro 2.0 (Naveed et al. 2021) the binding affinity between the multi-epitope vaccine and the TLR 2, 3 and 4 receptor was identified via molecular docking. ClusPro 2.0 worked in three stages at the same time. For instance, stiff body docking (Desta et al. 2020), clustering of lowest energy structure (Kozakov et al. 2013), and structural refinement by energy minimization. PIPER, a docking program based on the Fast Fourier Transform (FFT) correlation technique, is used in the rigid body docking step (Kozakov et al. 2017). The complex with the lowest energy score and the most effective docking was selected (Hossain et al. 2021).

Immune response simulation

The C-ImmSim, an agent-based approach to estimate the impact of antigen and foreign particles on immune function to determine the immune actions against the antigen and foreign particles (Oladipo et al. 2021). Three (3) injections were analyzed for 1050 simulations while leaving other parameters at default. The input file in FASTA format was the vaccine primary construct sequence. The model, known as C-IMMSIM, also includes a portion of (i) primary lymphatic organs where lymphocytes are formed and mature (most notably the red

bone marrow and the thymus gland), (ii) secondary lymphoid organs (for example, a lymph node), which filters lymph and is where naive B and T cells are exposed to antigens, and (iii) peripheral tissue that is reliant on the pathogen under study in this case (Castiglione et al. 2021). The operation of this tool is based on machine learning methods for subunit vaccination and binding interactions, as well as a position-specific score matrix. (Fadaka et al. 2021).

Molecular dynamics simulation

Molecular dynamics simulation is a well-known tool for analyzing biological systems at the molecular level. It was formerly used to test the stability of various protein complexes (Id et al. 2021a, b). CABSflex a tool (Jamroz et al. 2013) was used to predict the protein–ligand stability. CABSflex is a single-protein model-based integrated rapid technique for predicting protein structural variations and flexibility (Kuriata et al. 2018).The simulation length for cabs-flex has been tuned to get the highest feasible convergence with 10-ns molecular dynamics simulations (Jamroz et al. 2014). The PDB file of the vaccine-TLR complex was used as the input file.

Disulfide engineering

Disulfide by Design 2 v12.2 was used to do the vaccination protein disulfide engineering (Sarkar et al. 2020). DbD2 is a platform-agnostic web-based framework that improves accessibility, visualization, and analytical capabilities, (Hossain et al. 2021), implements the original design method again, makes a number of additions, and vastly enhances the web-based interface. (Craig and Dombkowski 2013). The server identifies which regions within a protein structure have the greatest chance of forming disulfide bonds. The intra-chain, inter-chain, and C for glycine residue were chosen during the prediction. The χ^3 and Cα–Cβ-Sγ angles were retained at 87 or +97 5 degrees and 114.610 degrees, respectively (Sarkar et al. 2020). The Fig. 2 below shows the overview of the entire procedure used in the process.
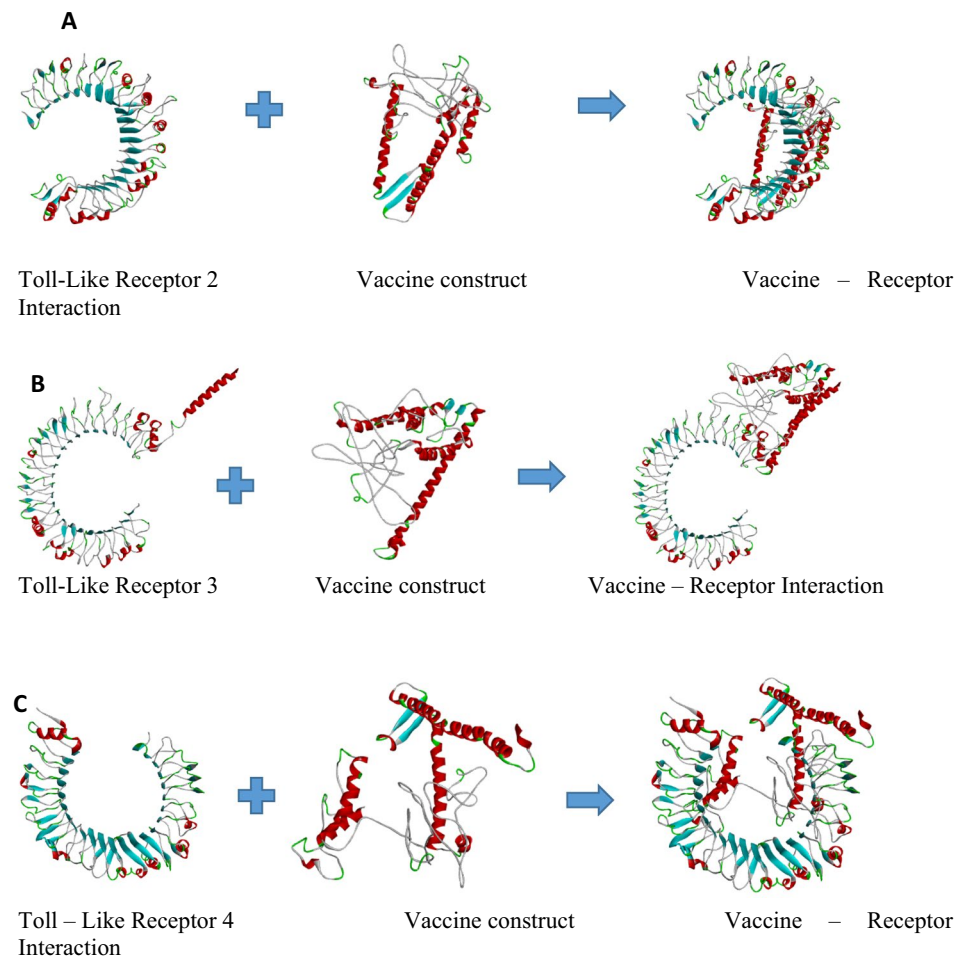
Result

Sequence retrieval and antigenicity prediction

Representative SARs-CoV-2 sequences were retrieved from Global Initiative on Sharing All Influenza Data (GISAID) Database, of which, a minimum of 12 sequences was downloaded for each of the known variants from across the continents of the world. Of the total number of downloaded sequences, 22.87% (113) were downloaded from Europe, 19.84% (98) from North America, 19.03%

Fig. 9 Molecular docking of the vaccine construct with TLR showing: **A** Docked complex of vaccine—TLR-2, **B** Docked complex of vaccine -TLR 3, and **C** Docked complex of vaccine—TLR-4

A

Toll-Like Receptor 2 Interaction   Vaccine construct   Vaccine – Receptor

B

Toll-Like Receptor 3   Vaccine construct   Vaccine – Receptor Interaction

C

Toll – Like Receptor 4 Interaction   Vaccine construct   Vaccine – Receptor

(94) from Asia, 13.56% (67) from Africa, 13.16% (65) from South America and 11.54% (57) from Oceania.

When these sequences were subjected to antigenicity test using Vaxijen 2.0 server at 0.5 threshold, only 60.73% (300) was found to be antigenic. The antigenic sequences were then subjected to further predictions and analyses.

Prediction and assessment of cytotoxic T lymphocytes (CTL) epitopes

From the sequences, more than 15,000 epitopes were predicted using NetCTL 1.2 server. Out of these enormous epitopes, only 125 were selected for further assessment based on the significant of the binding affinities they show towards MHC-1. Using ToxinPred and AllerTOP 2.0 servers respectively, 45 epitopes were found to be non-toxic while 19 were found to be non-allergen (Zhao et al. 2017). When these 19 immunogenic, non-toxic and non-allergenic epitopes were subjected to antigenicity test using Vaxijen 2.0 server at a threshold of 0.5, only 8 epitopes were found to be successful and were employed for vaccine construction. These 8 epitopes, additionally, were predicted to be

interferon-gamma (IFN-γ) inducing epitopes by the IFN-γ epitope server. IFN-gammas are known to evade intracellular pathogens and act as cytokines (Bibi et al. 2021).

Prediction and assessment of helper T lymphocytes epitopes

HTL epitopes are those high-binding MHC-II epitopes for human alleles HLA DR, predicted with the IEDB MHC-II web server. Furthermore, the predicted epitopes were subjected to toxicity, allergenicity and antigenicity tests, using appropriate servers respectively, a total of 2 epitopes were found to be suitable for and were employed in the construction of the vaccine. These selected epitopes were chosen on the bases that they are non-toxic, non-allergenic, and are antigenic.
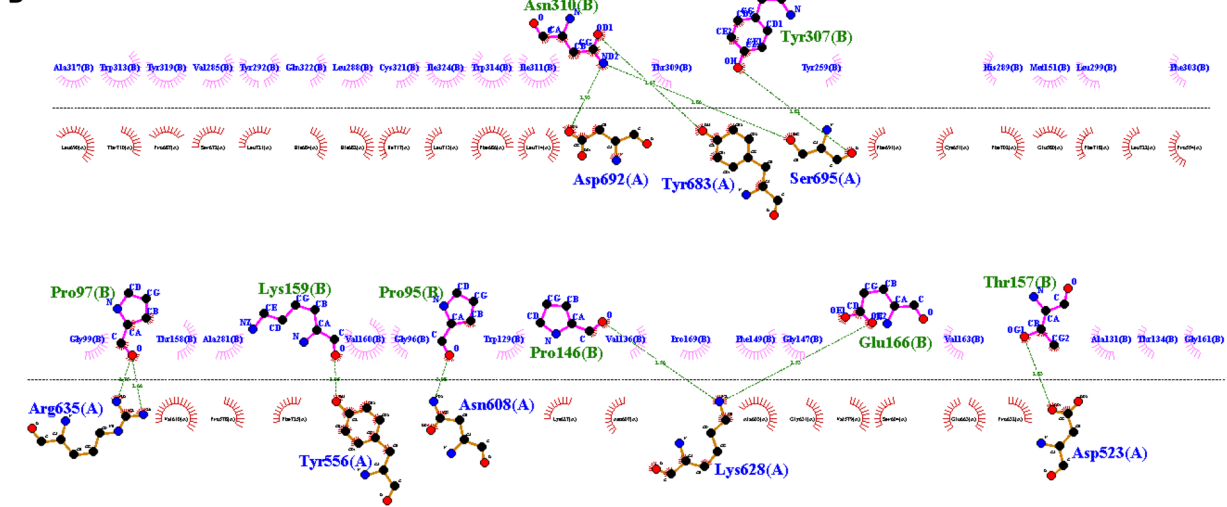
Prediction and assessment of B cell epitopes

In predicting the B cell epitopes, ABCPred server was used. The predicted B cell epitopes with binding score > 0.8 which
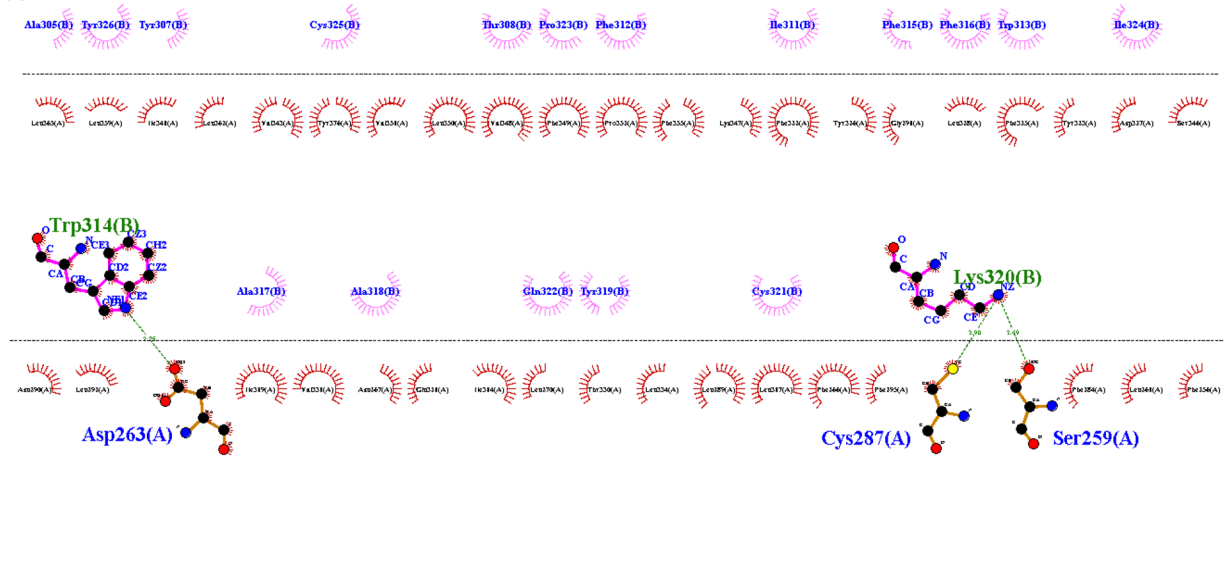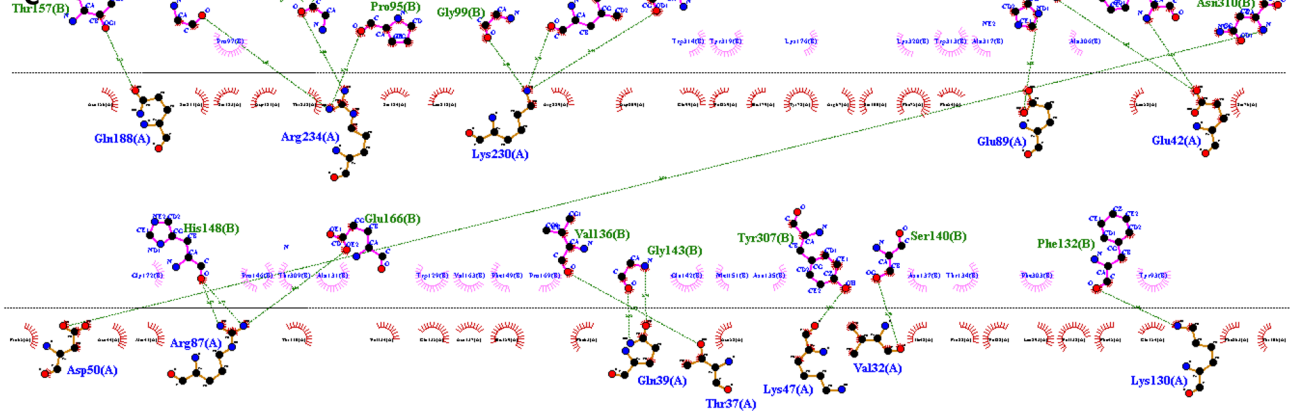
◀**Fig. 10** Vaccine—TLR interaction Visualization showing **A** Amino acid interaction between the vaccine and TLR-2 with the vaccine amino acids labeled green and receptor amino acids labeled blue. **B** Amino acid interaction between vaccine and TLR-3 with the vaccine amino acids labeled green and receptor amino acids labeled blue. **C** Amino acid interaction between vaccine and TLR-4 with the vaccine amino acids labeled green and receptor amino acids labeled blue

are antigenic, non-allergenic and non-toxic were selected for the vaccine construction. 5 epitopes were selected in total.

Toxicity, allergenicity, antigenicity, and solubility prediction

The antigenic score produced by using the Vaxijen 2.0 server to predict construct antigenicity is 0.6568, whereas the AntigenPro tool of the scratch protein server yields 0.939710, indicating that the construct is antigenic (Hossain et al. 2021). According to AllerTOP's allergenicity assessment, the vaccine design is non-allergenic. The ToxinPred server also anticipated that the recombinant vaccination design would be non-toxic. The construct's solubility, as determined by Protein-Sol, is 0.378, and its pI prediction is 9.900, which is within the range predicted by Expasy Protparam (9.60).

Physicochemical properties

The target protein construct employed in this investigation has an instability index score of 29.67, according to the ExPASy tool in ProtParam used for physicochemical prediction. This instability index indicated that the protein is stable (a score of > 40 indicates instability) (Hossain et al. 2021).The predicted aliphatic index score is 65.58, a protein with a high aliphatic index is thermostable over a wide temperature range, implying that the construct is stable at high temperatures (Pyasi et al. 2021). A GRAVY result of 0.120 was anticipated as a positive grand average of hydropathicity. This score indicates that the construct is hydrophobic, meaning that it has a low proclivity to interact with the surrounding water environment (negative value suggests high hydrophilicity) (Adhikari et al. 2018). The 326 amino acids in the intended design have a molecular weight of 35,667.15 kDa. The projected half-life for mammalian reticulocytes was 30 h, yeast > 20 h, and E. coli > 10 h.

Secondary structure prediction

The vaccine construct's secondary structure was predicted using two servers. Figure 3 shows the results of the secondary structure of the B-Cell construct as predicted by the SOPMA server. The alpha helix, beta-turn, extended strand, and random coil values for the B-Cell design were 26.69 percent, 4.60 percent, 28.53 percent, and 40.18 percent, respectively. The alpha helix, beta-turn,

signal peptide, and random coil are likewise predicted by PSIPRED V4.0, as illustrated in Fig. 4. Based on the results from the two servers, the vaccination had the highest percentage of amino acids in the coil shape. PSIPRED additionally predicts tiny nonpolar, hydrophobic, polar, and aromatic plus cysteine molecules, in addition to the aforementioned factors. Small non-polar amino acid has the highest % in the build of all the parameters anticipated (Fig. 5).

3D structure prediction

Scratch Protein, projected the vaccine construct's 3D structure. Figure 6 shows the final result for the tertiary structure. Pymol was used to examine the structure subsequently.

3D-structure refinement and validation

The 3D structure was polished further with the 3Drefine and GalaxyRefine tools, yielding 5 refined structures (Fig. 7A). Model 2 was chosen because it had a higher Rama preferred region (96.3) and generally acceptable GDT-HA (0.9095), RMSD (0.522), and MolProbity (1.945) scores (Table 1). In addition, the model's -1.71 Z-score was predicted by the ProSA-web z-score plot for the 3D structure for B Cell Construct (Figs. 7B and C). These findings suggest that the 3D-modeled structure is trustworthy. The vaccine subunit's Ramachandran plot revealed that 99.630 percent of the amino acid residues are in the high favored observation (Fig. 7D). Furthermore, the Ramachandran plot analysis projected that only 0.370 percent of the residues would be discovered in the preferred observation, whereas none would be found in the questionable observation (0.000 percent).

Codon optimization of designed vaccine peptide for expression analysis

The optimized codon was 157 nucleotides long, with a codon adaptation index (CAI) of 0.117 and a GC content of 45.79 percent before adaptation, a CAI of 1.0, and a GC content of 50.47 percent after adaptation for the pasted sequence. Because GC content ranges between 30 and 70% and is considered optimal, these values imply persistent vector expression in E. coli, and CAI values more than 0.8 are considered good for expression in the target organism.

Conformational B-cell epitope prediction

B cell epitope prediction based on conformation the discontinuous B-cell epitopes were identified using ElliPro at IEDB. The vaccine protein's PDB structure was utilized as an input for predicting conformational epitopes in the vaccine construct's tertiary structure. On the three-dimensional

A Antigen and immunoglobulins. Antibodies are sub-divided per isotype.
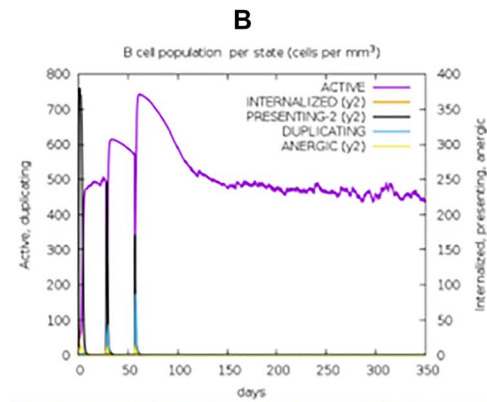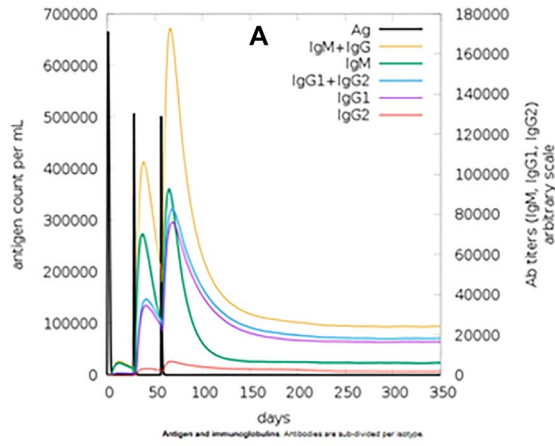
B **B lymphocytes** population per entity-state (i.e., showing counts for active, presenting on class-II, internalized the Ag, duplicating and anergic.
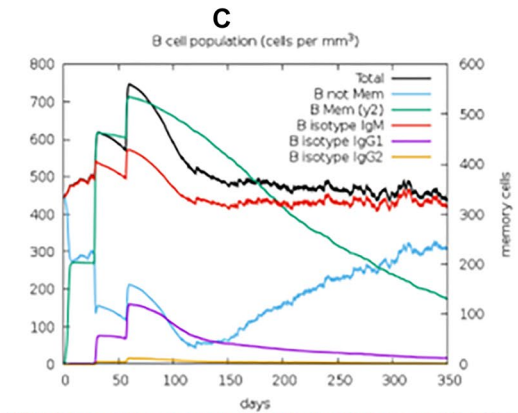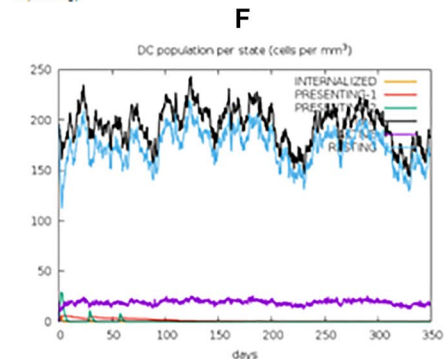
C **B lymphocytes**: total count, memory cells, and sub-divided in isotypes IgM, IgG1 and IgG2.
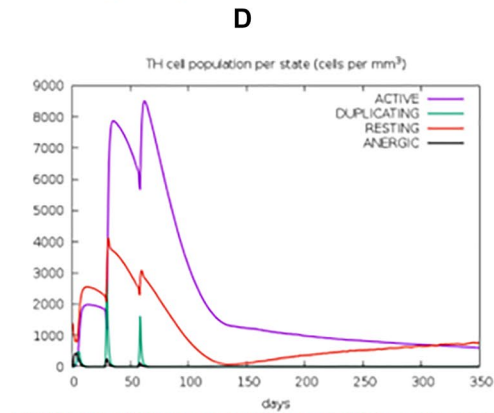
D **CD4 T-helper lymphocytes** count sub-divided per entity-state (i.e., active, resting, anergic and duplicating).
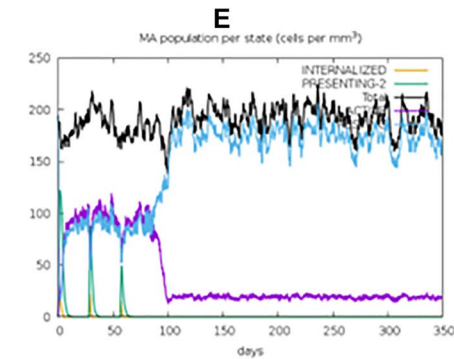
E **Macrophages**. Total count, internalized, presenting on MHC class-II, active and resting macrophages.

F **Dendritic cells**. DC can present antigenic peptides on both MHC class-I and class-II molecules. The curves show the total number broken down to active, resting, internalized and presenting the ag.
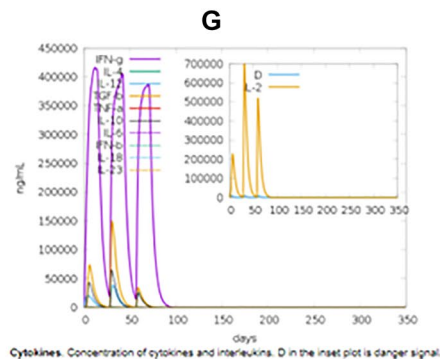
G **Cytokines**. Concentration of cytokines and interleukins. D in the inset plot is danger signal.

◄**Fig. 11** Immune simulation response as performed by C-SIM IMM for projected vaccine candidate **A** Immunoglobulins production in response to antigen with IgM+IgG labeled orange being the highest production and the antigen labeled black. **B** B–lymphocyte production showing its active state per day colored purple. **C** B-lymphocyte total count showing the production of memory cells per day of response colored green. **D** CD4 T-helper lymphocyte count showing the most active state being at the peak colored purple. **E** Macrophages total count. **F** Dendritic cells count and **G** Cytokines response

structure of the vaccine construct, ElliPro found eleven epitopes, with the maximum predicted score of 0.872 and a residual position of 301–326 (Table 2), and the structure of the preferred model is illustrated in Fig. 8.
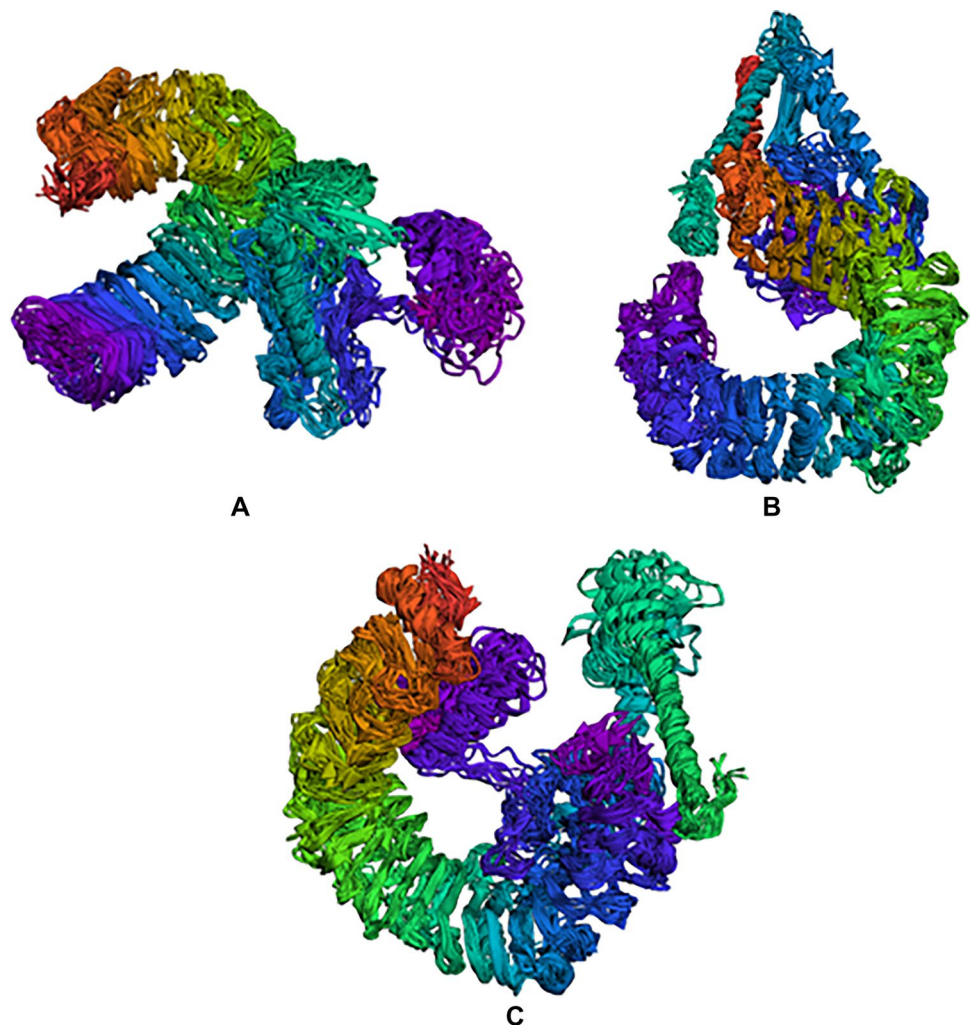
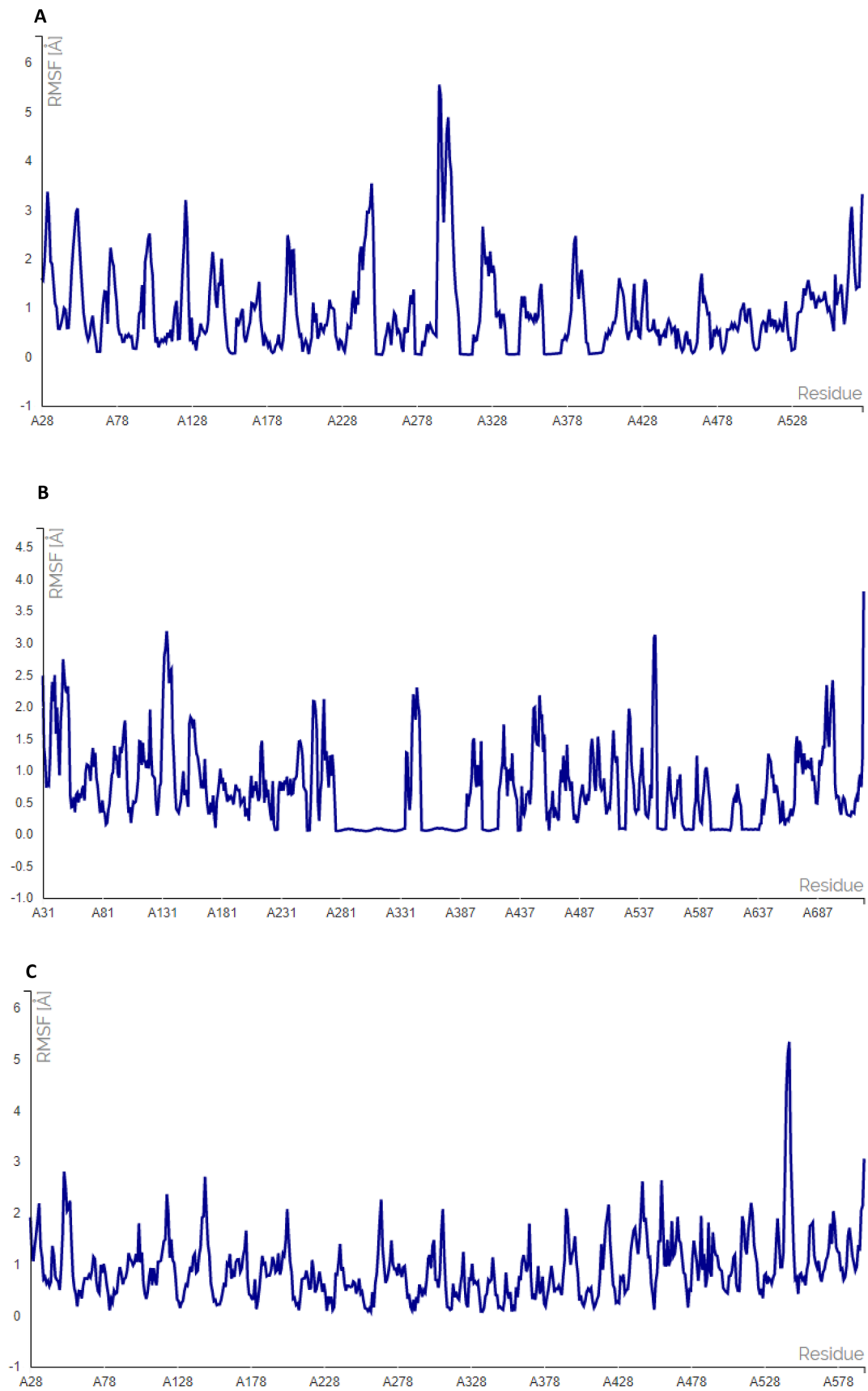## Molecular docking of vaccine construct with cleaned TLR 2, 3, and 4 receptors

Using the TLR as the receptor and the vaccine 3D construct as the ligand, the response of the novel vaccine design with different human TLRs (TLR2, TLR3, TLR 4) was investigated. Cluspro's output looked at 30 possible model outputs for each receptor-ligand reaction. The best model for the balanced coefficient was chosen based on binding energy weight. Figure 9 depicts the models that were chosen. To properly comprehend the connection between the vaccination candidates in the pocket of the TLRs. The toll-like receptor and vaccine amino acid interaction was extracted and visualized using the visualization tool LigPlot., as shown in Fig. 10.

Two (2) amino acids from the vaccine design (Trp314, Lys32) were discovered to bind to the pocket of TLR-2 with the amino acids Asp263, Cys287, and Ser 259. Additionally, eight (8) amino acids (Asn310, Tyr307, Pro97, Lys159, Pro95, Pro146, Glu166, Thr157) linked to TLR-3's pocket with Asp692, Tyr683, Ser695, Arg635, Tyr556, Asn608, Lys628, Asp523). Finally, sixteen (16) amino acids linked to the pocket of TLR-4 with amino acids Gln188, Arg234, Lys230, Glu89, Glu42, Asp50, Arg87, Gln39, Thr37, Lys47, Val32, Lys130.

Fig. 12 The spin direction of vaccine-Receptor complex **A** Vaccine-TLR-2 complex interaction **B** Vaccine-TLR-3 complex interaction **C** Vaccine-TLR-3 complex interaction

A

B

C

Immune response simulation

With three (3) injections, the C-IMM SIM server was used to precisely analyze the immunological response of the novel vaccine candidate as well as the increased half-life. As demonstrated in Fig. 11, the antigen and immunoglobulin responses, B-lymphocyte count, CD4 T-helper lymphocyte count, macrophage total count, dendritic cell count, and cytokine response were all shown.

Molecular dynamics simulation

The vaccine-receptor complex fluctuation and stability were investigated using molecular dynamics modeling. Figure 12 depicts the complexes' spin direction model as revealed by
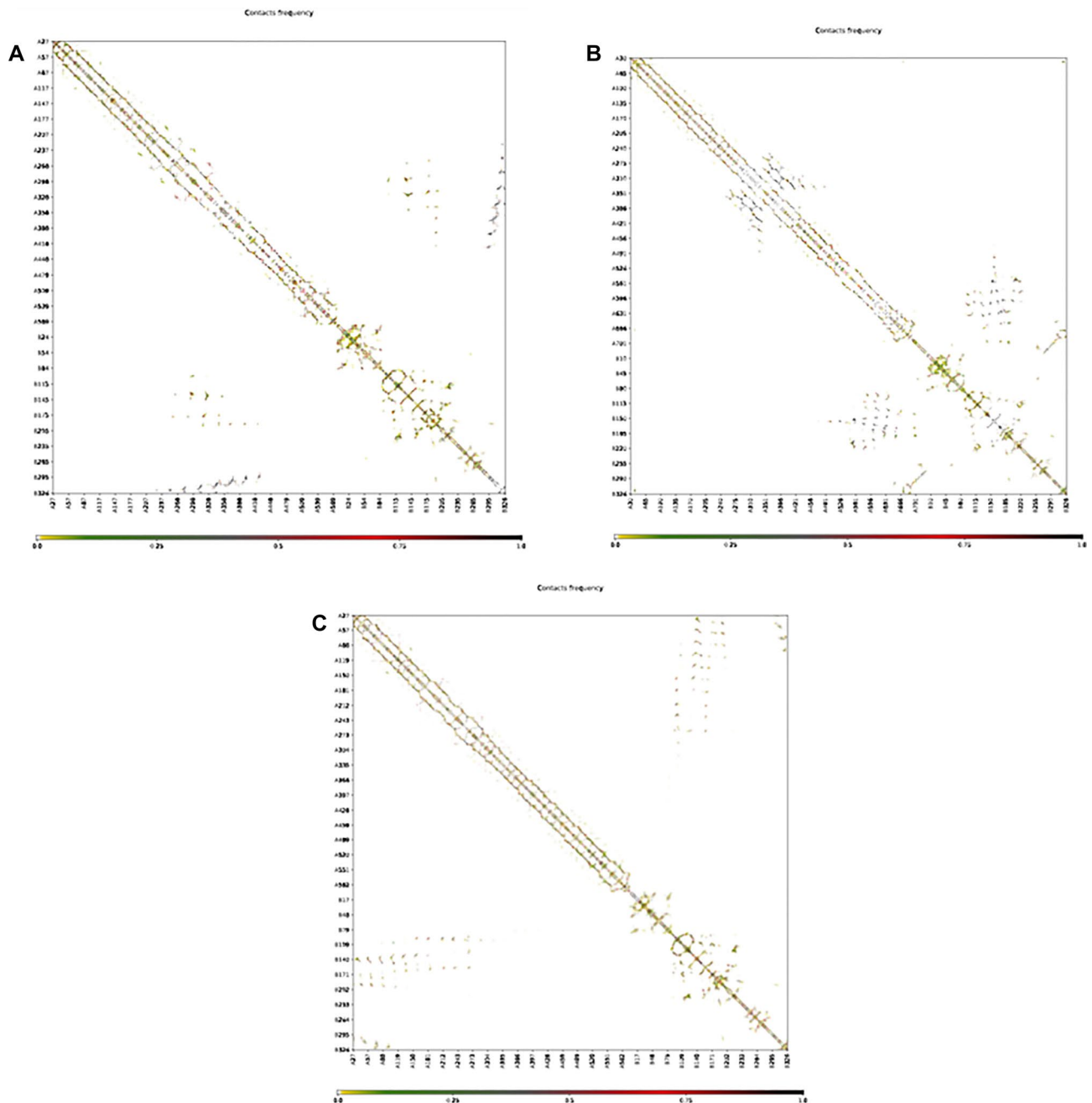
Fig. 14 Molecular dynamics contact map result **A** Contact map of vaccine-TLR-2 complex **B** Contact map of vaccine-TLR-3 complex **C** Contact map of the vaccine-TLR-4 complex

the server. The fluctuation plot (Fig. 13) and contact maps were also displayed on the server (Fig. 14).

Disulfide engineering

We examined the model with an energy score of less than 2.2 and a Chi3 value between + 97 and 87 based on the findings returned after exposing the vaccine construct to de-novo disulfide engineering. Taking the criteria into account. At positions 55, 89, 116, 191, 266, and 267, six (6) amino acids have chi3 values in this range, while just one amino acid has an energy score less than 2.2. Table 3 below shows the results in greater detail.

Discussion

The pressing global need for a SARS-CoV-2 vaccine to combat the COVID-19 pandemic stems from the absence of an effective treatment for this highly transmissible and severe respiratory syndrome (Anasir and Poh 2016). Vaccination is universally acknowledged as the most potent defense against infectious diseases, and COVID-19 is no exception. The concept behind peptide vaccines revolves around the identification and synthetic production of immunodominant B-cell and T-cell epitopes capable of eliciting specific immune responses (Patronov and Doytchinova 2013). However, with the emergence of various SARS-CoV-2 variants, currently available vaccines have proven inadequate against these new strains, necessitating the development of a universal vaccine capable of addressing all identified variants of concern (VOC) and variants of interest (VOI) as classified by the World Health Organization (WHO). Consequently, this study focused on designing a comprehensive multi-epitope peptide subunit vaccine, employing the spike glycoprotein as the primary reference point (Oladipo et al. 2021).

Bioinformatics analysis enables the efficient and accelerated prediction of potent epitopes, revolutionizing vaccine design compared to traditional approaches. The spike protein, primarily located on the outer surface of the virion, emerges as a promising target for B-cell epitope exploration due to its significant exposure. Previous investigations have revealed the ability of spike proteins from MERS-CoV and SARS-CoV to elicit robust immune responses. Baruah and Bose's study further corroborated the potential of spike protein epitopes as favorable candidates for the development of a SARS-CoV-2 vaccine (Chen et al. 2020a, b). We retrieved 494 sequences as representatives from six continents on Global Initiative on Sharing all Influenza Data (GISAID) and took into account all Variants of Concern (VOC) and Variants of Interest (VOI) as captured by WHO (Tracking the SARs-CoV-2 variants) (Oladipo et al. 2022a; Oude et al. 2023). The antigenic properties of the sequences were

predicted at a viral threshold of 0.5, sequences which falls below this threshold were considered to be non-antigenic and are therefore discarded, afterwards the B-cells, CTL, and HTL epitopes were predicted (Singh et al. 2020).

The prediction of B-cells and T-cells, which play a crucial role in mediating the adaptive immune response, holds significant importance in combatting various infectious diseases, including SARS-CoV-2. This intricate process involves the differentiation of these cells into plasma cells, responsible for synthesizing antibodies that actively combat foreign substances within the body (Fadaka et al. 2021). The selection of helper T-lymphocytes (HTL) was based on the recommendations from the Immune Epitope Database (IEDB) and encompassed HLA-DRB1*03:01, HLA-DRB1*07:01, HLA-DRB1*15:01, HLA-DRB3*01:01, HLA-DRB3*02:02, HLA-DRB4*01:01, and HLA-DRB5*01:01 (Vita et al. 2015).

The comprehensive evaluation of predicted T lymphocyte cells, encompassing both helper T-lymphocyte (HTL) and cytotoxic T-lymphocyte (CTL) epitopes, aimed to determine their coverage across a diverse global population. The CTL epitopes exhibited an extensive reach, covering 68.87% of the entire world population, with notable coverage exceeding 70% in regions such as West Africa, West Indies, South Asia, and Europe (Oladipo et al. 2022b). Moreover, significant coverage exceeding 60% was observed in areas including North Africa, North America, Northeast Asia, South Africa, and Southwest Asia. Similarly, the predicted HTL epitopes exhibited broad coverage, encompassing 69.46% of the world population, with over 70% coverage in Europe and North America, as well as 60% coverage in East Asia, North Africa, South Asia, and the West Indies. Comparatively, the obtained results (Samad et al. 2022) exceeded the threshold value, highlighting the potential of these epitopes as strong vaccine candidates capable of inducing immune responses across multiple Human Leukocyte Alleles.

The final construct was meticulously crafted by integrating the cytotoxic T-lymphocyte (CTL), helper T-lymphocyte (HTL) and B-cell epitopes, along with their requisite linkers. To enhance the vaccine's stability and longevity, the adjuvant (β-Defensin) was conjugated to the CTL epitope via the EAAAK linker. This strategic design facilitates the generation of robust cellular and humoral immune responses specific to the target antigens, ensuring a comprehensive immune defense (Samad et al. 2022). Additionally, the construct incorporates other essential linkers such as GPGPGP and AAY (Ala-Ala-Tyr), serving as intra HTL, B-cell, and CTL epitope connectors. These linkers not only enhance the immune response against the pathogen but also prevent any potential immunogenicity arising from junctional regions, thereby optimizing the vaccine's efficacy.

The universal vaccine candidate's physiochemical properties were predicted, bestowing some useful characteristics

Table 3 Possible disulfide bond between the vaccine construct amino acids

Residue 1				Residue 2				Bond		
Chain	Seq #	AA	Struct	Chain	Seq #	AA	Struct	χ^3	kcal/mol	$\sum$ B-factor
	55	ILE			63	LYS		− 89.28	0.46	0.00
	73	LYS			76	SER		+ 100.04	2.20	0.00
	89	ILE			110	THR		+ 93.10	6.07	0.00
	116	VAL			122	GLY		− 87.10	3.18	0.00
	191	PHE			194	PRO		− 86.72	5.29	0.00
	266	VAL			279	PRO		− 82.69	5.68	0.00
	267	THR			270	ALA		− 88.53	3.04	0.00

of a prospective vaccine, such as the peptide's 326 amino acid residues, which are suited for a vaccine build and have a molecular weight of 35,667.15 Da. The hypothesized PI value-based vaccination nature was discovered to be somewhat basic. According to the tools' predicted stability index, the protein is predicted to be stable after synthesis. However, the GRAVY value (29.67) and aliphatic index (65.58) showed that the vaccine was both thermostable and hydrophobic. The likelihood of designating this vaccine as a viable candidate is greatly increased by the positive physicochemical properties anticipated for the vaccine and by all of the scores on various metrics (Samad et al. 2022).

The determination of the peptide sequence's shelf life involved estimating its stability under various conditions. The analysis revealed that the potential universal vaccine candidate exhibited a longevity of approximately 30 h in mammalian reticulocytes (in vitro), while demonstrating shorter durations of less than 20 h in yeast and less than 10 h in Escherichia coli (in vivo). These estimations were derived using bioinformatics tools, specifically the protparam, and aligned with the findings of previous studies (Samad et al. 2022). Furthermore, a solubility test was conducted, indicating that the vaccine candidate displayed lower solubility with a value of 0.378, falling below the threshold of 0.45.

The peptide designed as potent vaccine has been shown to have antigenic properties when analyzed with vaxijen at 0.6568 and it was also validated with Antigen Pro at a value of 0.939710 which also confirm it antigenic characteristics. Its toxicity analysis when verified with toxinpred indicate that the peptide is non-toxin molecule. For the vaccine construct to be a potent candidate, the construct was analyzed for its allergenicity capacity, which it is shown that the construct was non allergen as predicted by AllerTop 2.0 and confirmed by AllergenFP. The secondary characters of the structure were predicted using SOPMA server and justified with PSIPRED, these two bioinformatics tools indicate the secondary structures has an alpha helix, Beta turn, extended strand and random coil.

Following this result, the 3D structure as shown by scratch protein and was validated and refined.

Furthermore, to access it compatibility against Human Toll Like Receptors using insilico design. we docked the potential vaccine candidate against TLR 2, 3, 4 (Singh et al. 2020) and these also enables us to visualize the interactions between these molecules using the Ligplot. The human toll like receptor 4 shows better compatibility as similar works were carried out by Oladipo et al. (2021); Oladipo et al. (2020) and Samad et al. (2022). This interaction shows that there is probable sixteen amino acids to bind with therefore increasing it suitability.

Immune simulation, a technique that was carried out to indicate the possible effect on the of the vaccine activities on the immune cells on the host. With the administration of first (initial) dose, the second dose after 28 days and the booster dose after few months. Immune titre curves which show a parameter that approves it usage and could prove proper Immune response (Bassaganya-Riera and Hontecillas 2016). Molecular dynamics simulation between vaccine and the TLR 4 was done inorder to investigate the interactions and stability between these biomolecules. This process shown by Cabsfleck in the flunctation plots shown above and the procedure was carried out under 10 nansec. Which agrees Jain et al. (2014). Disulsifide engineering was done to access the energy bond between the amino acids residue and this conforms with Oladipo et al. (2021).

The application of immunoinformatics techniques entails scrutinizing SARs-CoV-2 variants, unveiling vaccine constructs with promising potential as a universal Covid-19 vaccine candidate. However, the claim necessitates validation through subsequent wet lab assessments.

Conclusion

The study examined the genomic sequence of a representative sample of circulating SARs-CoV-2 variants in an effort to identify a possible vaccine candidate for the dreadful

Covid 19 illness. These peptide constructions show promising properties in in silico assessments as a potential all-purpose vaccine and may go on to wet lab assessments for the creation of a targeted vaccination against Covid 19 infection.

Data availability The data used during the course of this project where publicly available biological data retrieved from GISAID (https://gisaid.org/), and would be provided upon request.

Declarations

Conflict of interest The authors declare that we have no known competing financial interest or personal relationships that could have appeared to influence the work reported in this paper.

References

Adhikari UK, Tayebi M, Rahman MM (2018) Immunoinformatics approach for epitope-based peptide vaccine design and active site prediction against polyprotein of emerging oropouche virus. J Immunol Res 2018:1–22

Ali NS, Ahmad HZ, Qasim M, Suleman M, Ali S, Khan A, Wei DQ (2022) Computational modeling of immune response triggering immunogenic peptide vaccine against the human papillomaviruses to induce immunity against Cervical Cancer. Viral Immunol 34(7):457–469. https://doi.org/10.1089/vim.2020.0306

Anasir MI, Poh CL (2016) structural vaccinolody for viral vaccine design. Front Microbiol 10:738. https://doi.org/10.3339/fmicb.2019.007387

Anderson RJ, Weng Z, Campbell RK, Jiang X (2005) Main-chain conformational tendencies of amino acids. Proteins 689:679–689. https://doi.org/10.1002/prot.20530

Bassaganya-Riera J, Hontecillas R (2016) Chapter 1 - Introduction to Computational Immunology. In J Bassaganya-Riera (ed), Computational Immunology (pp 1–8). Academic Press. https://doi.org/10.1016/B978-0-12-803697-6.00001-1

Bibi S, Ullah I, Zhu B, Adnan M, Liaqat R, Kong WB, Niu S (2021) In silico analysis of epitope-based vaccine candidate against tuberculosis using reverse vaccinology. Sci Rep 11(1):1249. https://doi.org/10.1038/s41598-020-80899-6. PMID: 33441913; PMCID: PMC7807040

Castiglione F, Deb D, Srivastava AP, Liò P (2021) From infection to immunity: understanding the response to SARS-CoV2 through in-silico modeling. Front Immunol 12:1–16. https://doi.org/10.3389/fimmu.2021.646972

Chandra SP (2012) Textbook of microbiology and immunology, 2nd edn. Elsevier Inc.

Chen H, Tang L, Yu X, Zhou J, Chang Y, Wu X (2020a) Bioinformatics analysis of epitope-based vaccine design against the novel SARS-CoV-2. Infect Dis Poverty 9:1–10

Chen Y, Liu Q, Guo D (2020b) Emerging coronaviruses: genome structure, replication and pathogenesis. J Med Virol 92(4):418–423. https://doi.org/10.1002/jmv.252681

Craig DB, Dombkowski AA (2013) Disulfide by Design 2.0: a web-based tool for disulfide engineering in proteins Disulfide by

Design 2.0: a web-based tool for disulfide engineering in proteins. BMC Bioinformatics 14:0–6

Desta IT, Porter KA, Xia B, Kozakov D, Vajda S, Desta IT, Porter KA, Xia B, Kozakov D, Vajda S (2020) Resource performance and its limits in rigid body protein- protein docking. Structure/folding Des 28(9):1071-1081.e3. https://doi.org/10.1016/j.str.2020.06.006

Doytchinova IA, Flower DR (2008) Bioinformatic approach for identifying parasite and fungal candidate subunit vaccines. Open Vaccine J. 1:22–26

Ebrahimi M, Javad S, Mohammad S, Ranjbar M, Tahvildar F (2019) Designing and modeling of multi - epitope proteins for diagnosis of toxocara canis infection. Int J Pept Res Ther. https://doi.org/10.1007/s10989-019-09940-1

Fadaka AO, Remaliah N, Sibuyi S, Martin DR, Goboza M, Klein A, Madiehe AM, Meyer M (2021) Immunoinformatics design of a novel epitope - based vaccine candidate against dengue virus. Sci Rep. https://doi.org/10.1038/s41598-021-99227-7

Hossain S, Imran M, Mizan S, Tayab A, Sultana A (2021) Informatics in medicine unlocked immunoinformatics approach to designing a multi-epitope vaccine against Saint Louis encephalitis virus. Inform Med Unlocked 22:100500. https://doi.org/10.1016/j.imu.2020.100500

Id OSC, Gray CM, Fagbayi TA, Chukwuanukwu C, Oyebanji VO, Bankole TT, Adewole RA, Daniel M (2021a) Immuno-informatics design of a multimeric epitope peptide based vaccine targeting. PLoS One. https://doi.org/10.1371/journal.pone.0248061

Id TK, Id RI, Jahan T, Id J, Mahmud A (2021b) Immunoinformatics and molecular dynamics approaches: next generation vaccine design against West Nile virus. Cdc. https://doi.org/10.1371/journal.pone.0253393

Jain A, Tripathi P, Shrotriya A, Chaudhary R, Singh A (2014) In silico analysis and modeling of putative T cell epitopes for vaccine design of Toscana virus. J 3 Biotech 5:497–503

Jamroz M, Kolinski A, Kmiecik S (2013) CABS-flex: server for fast simulation of protein structure fluctuations. Nucleic Acids Res 41(May):427–431. https://doi.org/10.1093/nar/gkt332

Jamroz M, Kolinski A, Kmiecik S (2014) arXiv : 1403.4954v1 [q-bio. BM] 19 Mar 2014 ensembles. https://doi.org/10.1093/bioinformatics/btu184

Kozakov D, Beglov D, Bohnuud T, Mottarella SE, Xia B, Hall DR, Vajda S (2013) How good is automated protein docking? Proteins. 81(12):2159–2166. https://doi.org/10.1002/prot.24403. Epub 2013 Oct 17. PMID: 23996272; PMCID: PMC3934018

Kozakov D, Hall DR, Xia B, Porter KA, Padhorny D, Yueh C, Beglov D, Vajda S (2017) The ClusPro web server for protein-protein docking. Nat Protoc 12(2):255–278. https://doi.org/10.1038/nprot.2016.169. Epub 2017 Jan 12. PMID: 28079879; PMCID: PMC5540229

Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol 35(6):1547–1549. https://doi.org/10.1093/molbev/msy096

Kuriata A, Gierut AM, Oleniecki T, Pawe- M, Kolinski A, Kurcinski M, Kmiecik S (2018) CABS-flex 20: a web server for fast simulations of flexibility of protein structures. Nucleic Acids Res 46:338–343. https://doi.org/10.1093/nar/gky356

Madlala T, Adeleke VT, Fatoba AJ, Okpeku M, Adeniyi AA, Adeleke MA (2021) Designing multiepitope - based vaccine against Eimeria from immune mapped protein 1 (IMP - 1) antigen using immunoinformatic approach. Sci Rep. https://doi.org/10.1038/s41598-021-97880-6

Mahmud S, Rafi O, Paul GK, Promi MM, Sharmin M, Shimu S, Biswas S, Emran TB, Dhama K (2021) Designing a multi - epitope vaccine candidate to combat MERS - CoV by employing an

immunoinformatics approach. Sci Rep. https://doi.org/10.1038/s41598-021-92176-1

Myers J (2021) COVID-19: What you need to know about the coronavirus pandemic on 20 October. World Economic Forum. https://www.weforum.org/agenda/2021/10/covid-19-coronavirus-pandemic-20-october-2021/

Nandakumar K (2020) COVID-19: emergence, spread, possible treatments, and global. Front Public Health. https://doi.org/10.3389/fpubh.2020.00216

Naveed M, Tehreem S, Arshad S, Aniqa S (2021) Design of a novel multiple epitope-based vaccine: an immunoinformatics approach to combat SARS-CoV-2 strains. J Infect Public Health 14(7):938–946. https://doi.org/10.1016/j.jiph.2021.04.010

No B, Grote A, Hiller K, Scheer M, Mu R, Hempel DC, Jahn D (2005) Jcat: a novel tool to adapt codon usage of a target gene to its potential expression host. Nucleic Acids Res 33:526–531. https://doi.org/10.1093/nar/gki376

Oladipo EK, Ajayi AF, Ariyo OE, Onile SO, Jimah EM, Ezediuno LO, Adebayo OI, Adebayo ET, Odeyemi AN, Oyeleke MO, Oyewole MP, Oguntomi AS, Akindiya OE, Olamoyegun BO, Aremu VO, Arowosaye AO, Aboderin DO, Bello HB, Senbadejo TY, Awoyelu EH, Oladipo AA, Oladipo BB, Ajayi LO, Majolagbe ON, Oyawoye OM, Oloke JK (2020) Exploration of surface glycoprotein to design multi-epitope vaccine for the prevention of Covid-19. Inform Med Unlocked 21:100438. https://doi.org/10.1016/j.imu.2020.100438. Epub 2020 Oct 4. PMID: 33043110; PMCID: PMC7533051.

Oladipo EK, Folorunsho A, Olugbenga A, Onile S, Elijah O, Jimah EM, Ezediuno LO, Adebayo OI, Adebayo ET, Odeyemi AN, Oyeleke MO, Precious M, Ayomide O, Oguntomi S, Elizabeth O (2021) Designing a conserved peptide - based subunit vaccine against SARS-CoV-2 using immunoinformatics approach. In Silico Pharmacol. https://doi.org/10.1007/s40203-020-00062-x

Oladipo EK, Adeniyi MO, Ogunlowo MT, Irewolede BA, Adekanola VO, Oluseyi GS, Omilola JA, Udoh AF, Olufemi SE, Adediran DA, Olonade A, Idowu UA, Kolawole OM, Oloke JK, Onyeaka H (2022a) Bioinformatics designing and molecular modelling of a universal mrna vaccine for SARS-CoV-2 infection. Vaccines 10(12):2107. https://doi.org/10.3390/vaccines10122107

Oladipo EK, Akindiya OE, Oluwasanya GJ, Akanbi GM, Olufemi SE, Adediran DA, Bamigboye FO, Aremu RO, Kolapo KT, Oluwasegun JA, Awobiyi HO, Jimah EM, Irewolede BA, Folakanmi EO, Olubodun OA, Akintibubo SA, Odunlami FD, Ojo TO, Akinro OP, Hezikiah OS, Babalola MO (2022b) Bioinformatics analysis of structural protein to approach a vaccine candidate against Vibrio cholerae infection. Immunogenetics 75(2):99–114. https://doi.org/10.1007/s00251-022-01282-5

Onile OS, Ojo GJ, Oyeyemi BF, Agbowuro GO, Fadahunsi AI (2020) Development of multiepitope subunit protein vaccines against *Toxoplasma gondii* using an immunoinformatics approach. NAR Genom Bioinform 2(3):1–11. https://doi.org/10.1093/nargab/lqaa048

Patronov A, Doytchinova I (2013) t cell epitope vaccine design by immunoinformatics. Open Biol 3(1):120139. https://doi.org/10.1098/rsob.120139

Ponomarenko J, Bui HH, Li W, Fusseder N, Bourne PE, Sette A, Peters B (2008) ElliPro: a new structure-based tool for the prediction of antibody epitopes. BMC Bioinformatics 9:1–8. https://doi.org/10.1186/1471-2105-9-514

Pharmacy Practice News (nd) COVID-19 Patients Don't Need "Just in Case" Antibiotics. Retrieved July 24, 2023, from https://www.pharmacypracticenews.com/Covid-19/Article/09-20/COVID-19-Patients-Don-t-Need-Just-in-Case-Antibiotics/59608

Pyasi S, Sharma V, Dipti K, Jonniya NA (2021) Immunoinformatics approach to design multi-epitope- subunit vaccine against bovine ephemeral fever disease. Vacacines 9:1–20

Rapin JR, Wiernsperger N (2010) Possible links between intestinal permeability and food processing: A potential therapeutic niche for glutamine. Clinics (Sao Paulo) 65(6):635–643. https://doi.org/10.1590/S1807-59322010000600012. PMID: 20613941; PMCID: PMC2898551.

Rawal K, Sinha R, Abbasi BA, Chaudhary A, Nath SK, Kumari P, Preeti P, Saraf D, Singh S, Singh P, Sood S, Subramani P, Dubey AK, Strych U (2021) Identification of vaccine targets in pathogens and design of a vaccine using computational approaches. Sci Rep 0123456789:1–25. https://doi.org/10.1038/s41598-021-96863-x

Rehman A, Ahmad S, Shahid F, Albutti A, Alwashmi ASS, Aljasir MA, Alhumeed N, Qasim M, Ashfaq UA, Tahir M (2021) Integrated core proteomics, subtractive proteomics, and immunoformatics investigation to unveil a potential multi-epitope vaccine against schistosomiasis. Vaccines 9:1–21

Resolution GA, Others (2015) Transforming our world: the 2030 Agenda for Sustainable Development. UN Doc. A/RES/70/1. https://www.itu.int/en/ITU-T/Workshops-and-Seminars/2022/0808/Documents/Shailendra%20Kumar%20Mishra%20.pdf

Ryan LK, Diamond G (2017) Modulation of human β -defensin-1 production by viruses. Viruses 1:1–10. https://doi.org/10.3390/v9060153

Samad A, Ahammad F, Nain Z, Alam R, Imon RR, Hasan M, Rahman MS (2022) Designing a multi-epitope vaccine against SARS-CoV-2: an immunoinformatics approach. J Biomol Struct Dyn 40(1):14–30. https://doi.org/10.1080/07391102.2020.1792347. Epub 2020 Jul 17. PMID: 32677533; PMCID: PMC7441805.

Sanjua R (2016) Mechanisms of viral mutation. Cell Mol Life Sci. https://doi.org/10.1007/s00018-016-2299-6

Sarkar B, Ullah A, Araf Y, Shahedur M (2020) Informatics in Medicine Unlocked Engineering a novel subunit vaccine against SARS-CoV-2 by exploring immunoinformatics approach. Inform Med Unlocked 21:100478. https://doi.org/10.1016/j.imu.2020.100478

Shey RA, Ghogomu SM, Esoh KK, Nebangwa ND, Shintouo CM, Nongley NF, Asa BF, Ngale FN, Vanhamme L, Souopgui J (2019) In-silico design of a multi-epitope vaccine candidate against onchocerciasis and related filarial diseases. Sci Rep. https://doi.org/10.1038/s41598-019-40833-x

Singh A, Thakur M, Sharma LK, Chandra K (2020) Designing a multi - epitope peptide based vaccine against SARS - CoV - 2. Sci Rep 0123456789:1–12. https://doi.org/10.1038/s41598-020-73371-y

Sun C, Wang Y, Liu G, Liu Z (2020) Role of the eye in transmitting human coronavirus: what we know and what we do not know. Front Public Health 8(April):1–7. https://doi.org/10.3389/fpubh.2020.00155

Oude Munnink, B. B., & Koopmans, M. (2023). Tracking SARS-CoV-2 variants and resources [Review of Tracking SARS-CoV-2 variants and resources]. Nature Methods, 20(4), 489–490. https://doi.org/10.1038/s41592-023-01833-y

Vita R, Overton JA, Greenbaum JA, Ponomarenko J, Clark JD, Cantrell JR, Peters B (2015) The immune epitope database (IEDB) 3.0. Nucleic Acids Res 43(D1):D405–D412. https://doi.org/10.1093/nar/gku938

Yang Z, Bogdan P, Nazarian S (2021) An in silico deep learning approach to multi - epitope vaccine design : a SARS - CoV - 2 case study. Sci Rep 11:1–21

Zhang W, Liu J, Niu YQ, Wang L, Hu X (2008) A Bayesian regression approach to the prediction of MHC-I binding affinity. Comput Methods Progr Biomed 92(1):1–7. https://doi.org/10.1016/j.cmpb.2008.05.002

Zhao JW, Yan M, Shi G, Zhang SL, Ming L (2017) In silico identification of cytotoxic T lymphocyte epitopes encoded by RD5 region of *Mycobacterium tuberculosis*. J Inf Dev Countries 11(10):806–810. https://doi.org/10.3855/jidc.7207

Zhou Z, Dang Y, Zhou M, Li L, Yu C, Fu J, Chen S, Liu Y (2016) Codon usage is an important determinant of gene expression levels largely through its effects on transcription. Proc Natl Acad Sci. https://doi.org/10.1073/pnas.1606724113

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.