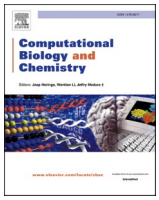
Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Computational Biology and Chemistry 101 (2022) 107754

Contents lists available at ScienceDirect

Computational Biology and Chemistry

journal homepage: www.elsevier.com/locate/cbac



Exploiting reverse vaccinology approach for the design of a multiepitope subunit vaccine against the major SARS-CoV-2 variants

Daniel Melo de Oliveira Campos [a], Maria Karolaynne da Silva [a], Emmanuel Duarte Barbosa [a], Chiuan Yee Leow [b], Umberto Laino Fulco [a], Jonas Ivan Nobre Oliveira [a,*]

[a] Department of Biophysics and Pharmacology, Bioscience Center, Federal University of Rio Grande do Norte, 59064-741, Natal/RN, Brazil
[b] Universiti Sains Malaysia, Penang, Malaysia

ARTICLE INFO

Keywords:
Immunoinformatics
Quantum mechanics/molecular mechanics calculations
Multi-epitope vaccine
SARS-CoV-2
Variants

ABSTRACT

The current COVID-19 pandemic, an infectious disease caused by the novel coronavirus (SARS-CoV-2), poses a threat to global health because of its high rate of spread and death. Currently, vaccination is the most effective method to prevent the spread of this disease. In the present study, we developed a novel multiepitope vaccine against SARS-CoV-2 containing Alpha (B.1.1.7), Beta (B.1.351), Gamma (P.1), Delta (B.1.617.2), and Omicron (BA.1) variants. To this end, we performed a robust immunoinformatics approach based on multiple epitopes of the four structural proteins of SARS-CoV-2 (S, M, N, and E) from 475 SARS-CoV-2 genomes sequenced from the regions with the highest number of registered cases, namely the United States, India, Brazil, France, Germany, and the United Kingdom. To investigate the best immunogenic epitopes for linear B cells, cytotoxic T lymphocytes (CTL), and helper T lymphocytes (HTL), we evaluated antigenicity, allergenicity, conservation, immunogenicity, toxicity, human population coverage, IFN-inducing, post-translational modifications, and physicochemical properties. The tertiary structure of a vaccine prototype was predicted, refined, and validated. Through docking experiments, we evaluated its molecular coupling to the key immune receptor Toll-Like Receptor 3 (TLR3). To improve the quality of docking calculations, quantum mechanics/molecular mechanics calculations (QM/MM) were used, with the QM part of the simulations performed using the density functional theory formalism (DFT). Cloning and codon optimization were performed for the successful expression of the vaccine in *E. coli*. Finally, we investigated the immunogenic properties and immune response of our SARS-CoV-2 multiepitope vaccine. The results of the simulations show that administering our prototype three times significantly increases the antibody response and decreases the amount of antigens. The proposed vaccine candidate should therefore be tested in clinical trials for its efficacy in neutralizing SARS-CoV-2.

1. Introduction

The SARS-CoV-2 pandemic is a global crisis that has yet to be resolved. Currently, the cumulative number of cases reported worldwide exceeds 396 million and the death toll is approximately 5.7 million Organization (2022). The mainstay of COVID-19 treatment is supportive care, but there is a high mortality rate, especially among the elderly and those with comorbidities. Intensive research is underway to find appropriate and effective therapies for the treatment and prevention of coronavirus infection, including safe and effective vaccines de Oliveira Campos et al. (2020); Campos et al. (2020a).

Vaccination is an effective means of improving public health by building adaptive immunity to a target pathogen Ehreth (2003). However, screening vaccine targets for clinical validation and production of a vaccine takes a long time. Advances in bioinformatics and next-generation sequencing technology, immunoinformatics, and reverse vaccinology can minimize the time required to screen antigens from protein sequences of pathogens and offer advantages in finding potential new vaccine targets Scarselli et al. (2005).

In recent months, a dizzying amount of information has emerged from numerous laboratories. Several vaccine candidates are currently in various stages of development, and a small number of vaccine candidates have reached clinical phases. As of April 26, 2022, the COVID − 19 vaccine candidate landscape reported 142 vaccines in clinical trials (31

* Corresponding author.
E-mail addresses: danielmelo.biomed@gmail.com (D.M.O. Campos), mkarolaynnesilva@gmail.com (M.K. Silva), emmanuel.baduarte@gmail.com (E.D. Barbosa), yee.leow@usm.my (C.Y. Leow), umbertofulco@gmail.com (U.L. Fulco), jonasivan@gmail.com.br, jonas.nobre@ufrn.br (J.I.N. Oliveira).

https://doi.org/10.1016/j.compbiolchem.2022.107754
Received 1 June 2022; Received in revised form 29 July 2022; Accepted 9 August 2022
Available online 18 August 2022

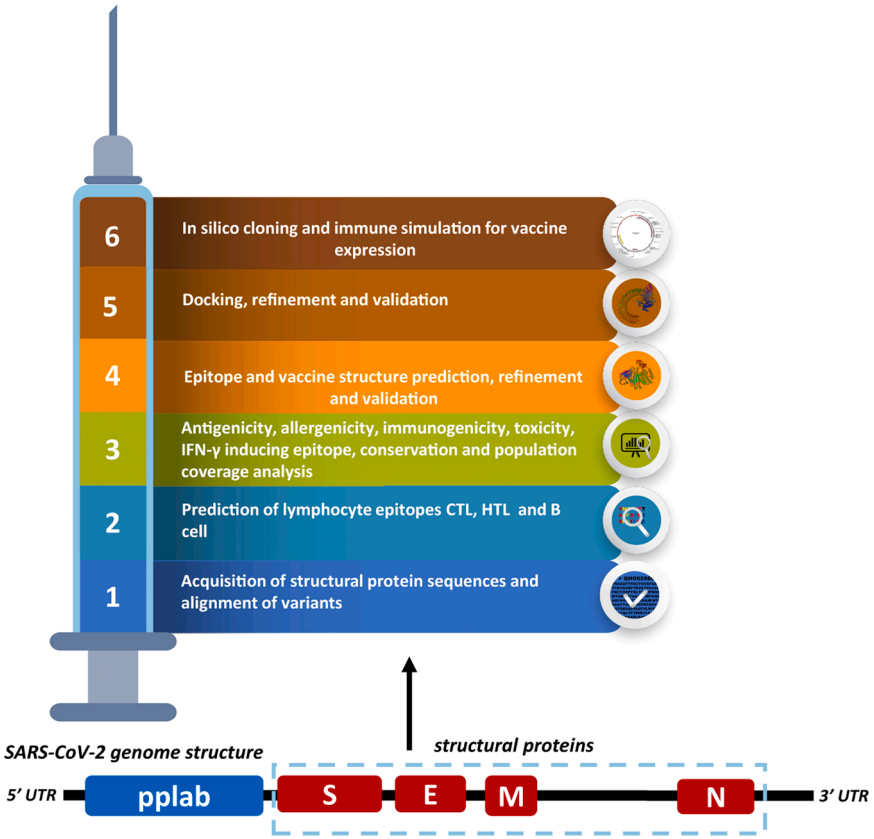


Fig. 1. Flowchart showing the stepwise methodology of predicting epitopes from structural proteins of the SARS-CoV-2.

in phase 3 and 10 in phase 4) and another 195 in preclinical phase WHO (2022).

The major vaccines being studied in clinical trials appear to be safe and effective means of preventing severe COVID-19, hospitalizations, and deaths, but the quality of evidence varies widely by vaccine. Questions remain regarding booster vaccination and waning immunity, duration of immunity, and heterologous vaccination Fiolet et al. (2021). In addition, the emergence of SARS-CoV-2 variants could jeopardize the global impact of mass vaccination campaigns, because it is already known that the natural tendency toward high mutation rates is the reason for the failure of most vaccines against viruses Zuin et al. (2021); de Oliveira Campos et al. (2022). After Alpha, Beta, Gamma, and Delta variants, the most recent variant of concern (VOC) is the Omicron (B.1.1.529), which has evolved due to the accumulation of a high number of mutations, particularly in the spike protein, raising concerns that it is capable of evading pre-existing immunity acquired through vaccination or natural infection. In fact, it is highly transmissible and has low susceptibility to neutralization by antibodies Doria-Rose et al. (2021); England (2021). Even after two doses of SARS-CoV-2 vaccine, neutralization of Omicron is far less than that of Delta or the parent viruses of SARS-CoV-2 Garcia-Beltran et al. (2021); Barda et al. (2021).

Conventional (biochemical, immunological, and microbiological) methods of vaccinology are inappropriate for this pandemic because antigen identification is time-consuming, culturing pathogens in laboratories is laborious, and costs are high María et al. (2017). Despite their success in many situations, these methods are unproductive for pathogens that cannot be cultured *in vitro* or for pathogens with highly variable antigen sequences such as HIV and influenza Stobart and Moore (2014). Alternatively, the methods of bioinformatics, vaccinogenomics, immunoinformatics, structural biology, and molecular simulations can be used for faster, more accurate, and less expensive vaccine development. They were first used to develop a vaccine against serogroup B *meningococci* and later against *Streptococcus pneumoniae*, *Staphylococcus aureus*, and *Chlamydia* Rappuoli (2001); Cui (2005). In recent years, these technologies have been successfully used in all phases of vaccine research, including preclinical, clinical, and postvaccine phases Soria-Guerra et al. (2015); Cui (2005); Soleymani et al. (2022); Elliott et al. (2008); Lennerz et al. (2014).

Here, we used consistent immunoinformatics and computational methods to [i] identify structural SARS-CoV-2 proteins from 475 genomes sequenced from regions with the most registered cases to cover all SARS-CoV-2 variants of concern. We also [ii] screened the best linear B-cell epitopes, CTL epitopes, and HTL epitopes from a robust research protocol with antigenicity, immunogenicity, allergenicity, toxicity, IFN-γ inducing, population coverage, and conservation filters. Furthermore, [iii] we designed, refined and validated the multi-epitope subunit vaccine; [iv] then we optimized the codon sequence and inserted it into a plasmid to ensure cloning and expression efficiency; [v] finally, we evaluated the synthesis of immunoglobulins, immune complexes, cytokines, and interleukin and thus the consistency of the immune response elicited by the prototype vaccine.

2. Methodology

The flow chart of the methodology used in this study is shown graphically in Fig. 1. Recently, our research group validated similar immunoinformatics and molecular modeling approaches in the construction of a multiepitope vaccine against Mayaro virus Silva et al. (2021); da Silva et al. (2022).

2.1. Acquisition of protein sequences

Initially, the primary sequences of SARS-CoV-2 structural proteins (S, M, N and E) of the Alpha (B.1.1.7), Beta (B.1.351), Gamma (P.1), Delta (B.1.617.2), and Omicron (BA.1) variants were extracted using the filters from the Virus Pathogen Resource (ViPR) database: severe acute

respiratory syndrome coronavirus 2; human (host); complete genome only. The viruses have several strains worldwide due to their natural tendency to show high mutational rates, which is could reason for fail of most of the vaccines Garcia-Boronat et al. (2008). For this reason, based on the multiple sequence alignment of 475 strains sequenced in one of the five countries with the highest number of cases (United States, India, Brazil, France, Germany, and United Kingdom), the sequential variability sites and the conserved fragments of four structural proteins of SARS-CoV-2 were mapped by the Protein Variability Server (PVS) (http://imed.med.ucm.es/PVS/) and MUSCLE algorithm.

2.2. Prediction of T cell epitopes

NetMHC 4.0 server (http://www.cbs.dtu.dk/services/NetMHC/) was used to predict the MHC I binding promiscuous epitopes in the consensus sequences, including 47 HLA class I alleles. Binding parameters using artificial neural networks (ANNs) Nielsen et al. (2003). Then, the NetCTL server (http://www.cbs.dtu.dk/services/NetCTL/) was used to predict cytotoxic T lymphocytes and values related to c-terminal cleavage, TAP transporter affinity, and HLA I binding affinity Peters et al. (2003). Models with high predictive accuracy were proposed for neural networks trained on 9-mer data, such that epitopes of this length are identified with higher sensitivity and specificity Lundegaard et al. (2008a), (2008b). Finally, the data obtained from both platforms were stored in spreadsheets and compared to analyze only the epitopes that were present on both servers.

For consensus prediction of MHC II-restricted peptide epitopes, the NetMHCII tool (http://www.cbs.dtu.dk/services/NetMHCII/) was used. This tool has an allele-specific method that includes an individual predictor for each MHC molecule in the dataset. Thus, affinities for MHC molecules can be predicted, with classification into binders and non-binders Jensen et al. (2018). The NetMHCIIpan (http://www.cbs.dtu.DK/services/NetMHCIIpan/) has also been used as a guide for T cell peptide identification. It is based on a comprehensive dataset of $>$ 100, 000 quantitative peptide binding measurements from IEDB, including 36 HLA-DR, 27 HLA-DQ, 9 HLA-DP, and 8 mouse molecules MHC-II Andreatta et al. (2015). In addition to accurate identification of the binding core, this pan-specific method allows: quantification of the probability of multiple binding colors within a single antigenic peptide and assignment of reliability scores for each binding core prediction Andreatta and Nielsen (2018).

Predicted binders were selected using a percentile ranking method comparing predicted affinity to a collection of 400,000 random endogenous peptides and MHC binding affinity. IC_{50} values of $<$ 50 nM and a percentile rank of $<$ 0.5 were classified as high affinity (or strong binding), 50–500 nM as medium affinity, and 500–5000 nM as low affinity Nielsen et al. (2007), (2009)). According to Vita et al. (2015), no T-cell epitope has reached an IC_{50} value of $>$ 5000 nM. Because regression at this scale (nM) is more challenging, it was linearized using the equation $log50k = 1 - log(IC_{50})/log(50000)$. To avoid false positives, all values in this log-transformed binding affinity ($1 - log50k$) identify peptides greater than or equal to 0.7 Fleri et al. (2017).

2.3. Antigenicity prediction

The antigenic properties of the selected theoretical epitopes were analyzed using VaxiJen 2.0 (http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html), an alignment-independent method developed to overcome the limitations of alignment-dependent sequence similarity methods and based on automatic cross-covariance transformation (ACC) of protein sequences into uniform vectors of key amino acid properties. Vaxijen predicts antigens based on protein physicochemical properties with 70–89 % accuracy Doytchinova and Flower (2007b).

2.4. Allergenicity prediction

Allergen identification is a crucial factor in the development of the vaccine. Therefore, AllerTOP v.2.0 server (http://www.ddg-pharmfac.net/AllerTOP/) measures the allergenic properties of epitopes based on an alignment-independent method that has been optimized, cross-validated, and implemented. It uses an updated set of 2427 known allergens and 2427 non-allergenic proteins from widely used foods and non-immunogenic human proteins. Data processing detects 87 % of allergens and 91 % of non-allergens in the external test set Dimitrov et al. (2014).

2.5. Immunogenicity prediction

To predict epitopes capable of eliciting both humoral and cellular immune responses, we used the IEDB immunogenicity prediction tool (http://tools.iedb.org/immunogenicity/) using the following parameters: cutoff equal to zero and standard mask Vita et al. (2015). We assessed the immunogenicity of the antigens by summing the immunogenicity scores of all epitopes predicted to bind the MHC-I reference set of alleles in each antigenic region of the proteins. A high score indicates a greater likelihood of eliciting an immune response Calis et al. (2013).

2.6. Toxicity test

TonxinPred (http://crdd.osdd.net/raghava/toxinpred/) is based on an SVM model for classifying toxicity and non-toxicity based on various physicochemical properties of peptides. The database used in this method consists of 1805 toxic peptides ($\leq$35 residues) and 3593 non-toxic peptides Gupta et al. (2013).

2.7. Conservation analysis

The IEDB Conservancy tool (http://tools.iedb.org/conservancy) was used to assess the degree of conservation of epitopes within the protein sequences of all available genotypes of SARS-CoV-2 obtained with varying degrees of sequence identity Bui et al. (2007). The degree of conservation can be defined as the proportion of protein sequences in which the epitope is present at a given degree of identity. In this way, it would be possible to verify which of the epitopes is most conserved and thus could become a vaccine candidate.

2.8. Population coverage study

The major histocompatibility complex (MHC) in humans refers to a genetic region containing hundreds of genes, including human leukocyte antigen (HLA) genes. T cells recognize a complex between a specific major histocompatibility complex (MHC) molecule and a specific pathogen-derived epitope, so the binding of peptides to MHC molecules is allele-specific. The selection of multiple peptides from SARS-CoV-2 with different HLA binding specificities will allow for greater coverage of the patient population targeted by peptide-based vaccines or diagnostics. Knowing that the small set of alleles covers most of the population, the IEDB Population Coverage tool (http://tools.iedb.org/population/) calculates the proportion of individuals from a given area (or worldwide) who are geographically predicted to respond to a given set of epitopes with known MHC restrictions based on their human leukocyte antigen (HLA) alleles Bui et al. (2006).

2.9. IFN-γ inducing epitope prediction

T helper cells can be identified by the production of various cytokines such as interferon-γ (IFN-γ), which subsequently play a role in the activation of various immune cells such as macrophages and natural killer cells Luckheeram et al. (2012). IFN-γ inducibility assessment of predicted HTL epitopes can be performed using the hybrid prediction

approach and the Support Vector Machine (SVM) of IFNepitope (http://crdd.osdd.net/raghava/ifnepitope/).

2.10. Linear B-cell epitope prediction

B-cell epitopes are antigenic determinants and have the ability to elicit humoral immunity recognized by B-cell receptors or secreted antibodies of the immune system and represent the specific part of the antigen to which B lymphocytes bind. The linear B-cell epitopes of all proteins were predicted using BepiPred 2.0 from IEDB (http://tools.iedb.org/bcell/). The algorithm is based on a random forest algorithm trained on epitopes annotated from antibody-antigen protein structures Ponomarenko and Bourne (2007).

2.11. Construction of multi-epitope vaccine sequence

The CTL and HTL epitopes identified by the above immunoinformatic methods were used to carefully construct a vaccine sequence. These CTL and HTL epitopes were linked using AAY and GPGPG linkers Kadam et al. (2020); Mittal et al. (2020). In addition, β-defensin was introduced into the N-terminal end of the vaccine construct due to its adjuvant activity against viral infections Ling et al. (2017).

2.12. Homology modelling and tertiary structure refinement

The 3D building models were performed using Swissmodel server (https://swissmodel.expasy.org/), which allows homology modeling and the creation of 5 final models. At the same time, we also submit the primary sequence to the Robetta server. The Robetta server (https://robetta.bakerlab.org/) generates structural models using either comparative modeling or de novo structure prediction methods. Then the best models for each server were selected for the refinement step.

The 3D structures were refined using the GalaxyRefine2 tool (http://galaxy.seoklab.org/). This server is based on a refinement method that performs short molecular dynamics relaxations (MD) after repeated side-chain repack perturbations, allowing larger motions. Experimentalists have used it in functional protein modeling studies to improve the quality of model structures obtained with other prediction methods. A recent benchmark test of CASP (Critical Assessment of techniques for protein Structure Prediction) refinement targets showed that GalaxyRefine2 was successful in performing blind predictions Lee et al. (2019).

2.13. Tertiary structure validation

Validation of the 3D model of the vaccine construct was performed using ProSA-web, ERRAT, and Verify3D. ProSA-web (https://prosa.services.came.sbg.ac.at/prosa.php) is a freely available web server that is commonly used to validate the input 3D model. It assigns a quality score to the input structure if the score falls outside a range typical of native proteins and the structure is most likely to have errors. The ERRAT (http://services.mbi.ucla.edu/ERRAT/) and Verify3D (https://servicesn.mbi.ucla.edu/Verify3D/) servers were used to determine the nonbonded interactions within the structure and to determine the compatibility of an atomic model with its amino acid sequence by assigning the class of the structure based on its environment and comparing the results with suitable structures.

The validation analysis was also performed using the Molprobity server (http://molprobity.biochem.duke.edu/), which provides a broad, robust analysis of model quality at global and local levels for protein structures. It relies on the performance and sensitivity provided by optimized hydrogen placement and analysis of contacts between all atoms, complemented by covalent geometry and torsion angle criteria Williams et al. (2018). The central statistic of protein structure quality is represented by the "molprobity score", a combination of clash score, rotamer, and Ramachandran assessments in a single score normalized to the same scale as X-ray resolution Chen et al. (2010).

2.14. Physiochemical properties of the vaccine

ProtParam (http://web.expasy.org/protparam/) was used to predict the physicochemical properties of the vaccine candidate, including molecular weight, theoretical pI, instability index, amino acid composition, grand average of hydropathicity (GRAVY), atomic composition, estimated *in vitro* and *in vivo* half-life, aliphatic index, and extinction coefficient. A protein is considered stable if its value is below the threshold of 40, while the hydropathy index evaluates the probability that a region is hydrophobic (positive values) or hydrophilic (negative values) Gasteiger et al. (2005).

2.15. Prediction of post-translational modifications

For the prototype vaccine, post-translational modification analysis was performed using network prediction tools. NetOGlyc 4.0 (http://www.cbs.dtu.dk/services/NetOGlyc/) and NetNGlyc 1.0 (http://www.cbs.dtu.dk/services/NetNGlyc/) were used for glycosylation, which is one of the significant posttranslational modifications Steentoft et al. (2013). In addition, this analysis included the prediction of phosphorylation using the NetPhos 3.1 server (http://www.cbs.dtu.dk/services/NetPhos/) Blom et al. (2004). Finally, the CSS-Palm server (http://csspalm.biocuckoo.org/online.php) predicts palmitoylation of proteins, an essential posttranslational lipid modification of proteins Ren et al. (2008).

2.16. Codon optimization of final vaccine constructs

The Java Codon Adaptation Tool (JCat) https://www.jcat.de/was used for codon optimization and back translation and evaluated with the Codon Adaptation Tool to predict the appropriate expression in vector translation and cloning efficiency Grote et al. (2005) and to generate the vaccine cDNA sequence required for efficient expression in the *Escherichia coli* K-12 strain Grote et al. (2005). In addition, the optimized multiepitope vaccine sequence was inserted into the pET-28a (+) vector using the SnapGene tool.

2.17. Immune simulation of the vaccine

The C-ImmSim server https://kraken.iac.rm.cnr.it/C-IMMSIM/was used to characterize the actual immunogenic profiles and immune response of the vaccine prototype. C-ImmSim is an online server that uses agent-based modeling to characterize the immune response profile and immunogenicity of the chimeric peptides. The model is based on position-specific scoring matrices (PSSM) for peptide prediction derived from machine learning techniques for predicting immune interactions Rapin et al. (2010). Two injections of the target prophylactic SARS-CoV-2 vaccine profile were administered at different 4-week intervals. Our simulation lasted 1050 time steps (one time step of approximately 8 h), or nearly 12 months. All other simulation parameters were kept as default values Castiglione et al. (2012).

2.18. Molecular docking, QM/MM study and binding profile

The vaccine prototype must bind with target immune cell receptors, such as Toll-Like Receptor 3 (TLR3) to elicit an efficient immune response Perales-Linares and Navas-Martin (2013). Using PROPKA 3.1 software (https://github.com/jensengroup/propka/), the molecular structures of the ligand (vaccine prototype) and receptor (TLR3 - PDB ID: 1ziw) were first modified by adding charges (protonation or deprotonation) to the atoms and correcting the bonds. Since SARS-CoV-2 is present in the bloodstream, the physiological value (7.2–7.4) was used as the pH parameter. Atomic optimization of the hydrogen geometry was performed using CHARMm (Chemistry at

Harvard Molecular Mechanics), a force field parameterized from molecular dynamics simulation specifically for organic molecules to increase the accuracy of the calculations Vianna et al. (2019); Campos et al. (2020b); Bezerra et al. (2020).

To evaluate the interaction of the vaccine prototype with the molecule TLR, we performed a structure-based docking analysis on the PatchDock server (http://bioinfo3d.cs.tau.ac.il/PatchDock/), a molecular docking algorithm based on the principles of shape complementarity, *i.e.*, molecular shape representation, surface patch matching, and partitioning and scoring Schneidman-Duhovny et al. (2005). First, the surfaces of the receptor and ligand molecules were partitioned into patches that corresponded to the shape of the surface. These patches then fit into specific patterns that allowed visual discrimination of the puzzle pieces. Once the patches were identified, shape matching methods were used to determine the superpositions of these patches.

Subsequently, the vaccine TLR3 candidates were refined using FireDock server (http://bioinfo3d.cs.tau.ac.il/FireDock/) by the constrained restructuring of the side chains at the interface and Monte Carlo minimization of the binding scoring function. The refined candidates were ranked based on a binding score parameterized using atomic contact energy, attenuated van der Waals interactions, partial electrostatics, and additional binding free energy estimates Andrusier et al. (2007).

We used the combined quantum mechanics/molecular mechanics technique (QM/MM) to select the most relevant vaccine-TLR3 (ligand-receptor) complex from the docking calculations. The QM/MM methods have now established themselves as the most advanced computational methods for biomolecular systems. The rapidly growing number of publications using QM/MM techniques is impressive evidence that they have come of age since their pioneering beginnings some thirty years ago Lindorff-Larsen et al. (2010); Senn and Thiel (2009). The QM/MM formalism allows the partitioning of the total energy and, in particular, the interaction energy into different components, thus offering the possibility of analyzing the effects of the protein environment (down to individual wastes), especially in the presence of many electrostatic interactions Chung et al. (2015).

The QM/MM optimization was performed within the framework of the ONIOM multilayer method (our own Integrated Molecular Orbital and N-layer Molecular Mechanics), which is available in Gaussian code. It is a robust method that allows accurate *ab initio* calculations of the total energy of large complexes, such as biochemical systems, when these systems are divided into two or three layers. In this case, the TLR3 receptor was assigned to the MM layer, while the major amino acid residues of the vaccine were assigned to the QM layer. The popular B3LYP hybrid functional (Becke, three parameters, Lee-Yang-Parr) of exchange-correlation and basis set 6–311 G (d, p) were used to expand the electronic orbitals for the QM layer, and all amino acid residues within a radius of 6.0 $\AA$ from the centroid of the ligand were allowed to move during geometry optimization.

Finally, binding poses and parameters were examined using Discovery Studio Visualizer (https://discover.3ds.com/discovery-studio-visualizer-download/), LigPlot+ (https://www.ebi.ac.uk/thornton-srv/software/LigPlus/), PoseView (https://proteins.Plus/), and: intermolecular hydrogen bonds (carbon, conventional, and pi-donor H-bonds), electrostatic (salt bridge, attractive charges, pi-cation, pi-anion), hydrophobic (pi-pi stacked, pi-pi stacked, alkyl, pi-sigma, pi-alkyl), halogens (Cl, fluorine, Br, and I), miscellaneous (charge repulsion, steric unevenness, acceptor-acceptor collision.

3. Results and discussion

Vaccination is a successful and cost-effective public health prevention strategy to combat deadly infectious diseases worldwide Chabot et al. (2004). The peptide-based vaccines or multiepitope vaccines have been shown to be the better choice for safe vaccination because they use short, nonallergenic peptide fragments to stimulate extremely targeted immunoprotective responses Li et al. (2014). Several previous *in silico* studies demonstrate the efficacy and effects of prototype multiepitope vaccines Yazdani et al. (2020); Rakib et al. (2020); AlSaba et al. (2021); Silva et al. (2021). Moreover, vaccine development in this manner has already gained momentum, with immunogenic potential demonstrated through *in vivo* experiments Lei et al. (2019); Zhao et al. (2021) and clinical trials Elliott et al. (2008); Lennerz et al. (2014).

In a critical global scenario triggered by the SARS-CoV-2 virus in recent years, the vaccines in the III and IV phases are relatively effective in controlling the disease, at least in exacerbating severe cases. However, the virus are known to have high genomic mutation rates that provide protection for them, and this is one of the main causes of vaccine failure Laughlin et al. (2015). Having traced the identified mutations of the SARS-CoV-2 virus to the present day (April 2022), it is clear that the accumulation of spike protein mutations is responsible for the increased infectivity and severity of SARS-CoV-2 and potentially hinders vaccine efficacy Harvey et al. (2021). Indeed, the emergence of variants of SARS-CoV-2 could compromise the global impact of mass vaccination campaigns Fiolet et al. (2021).

Recently, Zang et al. (2022) reported the development of an mRNA vaccine candidate specific to the receptor binding domain (RBD) of the Omicron variant. Two doses efficiently induce neutralizing antibodies in mice. However, the antisera are effective only in the Omicron variant and not in the wild-type and Delta strains. Lee et al. (2022) developed the "hybrid" vaccine containing the RBD with all 16 point mutations shown in Omicron and Delta RBD, and a bivalent vaccine consisting of Omicron and Delta RBD-L at half dose. Taken together, their data indicate that the Omicron-specific mRNA vaccine can elicit a strong neutralizing antibody response against Omicron, but that the inclusion of epitopes from other variants may be required to elicit cross-protection.

Therefore, we used an extensive immunoinformatics and molecular modeling protocol to develop a prototype vaccine against all of the above variants, in this case consisting of unique mutant epitopes of each variant.

3.1. Sequence retrieval

Here, the amino acids sequences for four structural proteins (Spike glycoprotein - S, membrane glycoprotein - M, envelope protein - E and nucleoprotein - N) of SARS-CoV-2 variants (B.1.1.7 - Alpha, B.1.351 - Beta, P.1 - Gamma, B.1.617.2 - Delta, and BA.1 - Omicron) were retrieved from the Vipr database Waterhouse et al. (2009) and used to predict the B and T cell epitopes. Thus, we encompassed 475 proteomes of the major circulating SARS-CoV-2 variants, including 38 (16) reported non-synonymous mutations found in the protein spike (nucleocapsid phosphoprotein).

3.2. Cytotoxic T Lymphocytes (CTL)

The major histocompatibility complex class I (MHC-I) processing pathway involves the degradation of protein antigens by a constitutive proteasome or immunoproteasome, resulting in the production of antigenic peptides. The Antigen Processing Protein-Associated (TAP) transporter transports peptides from the cytosol to the endoplasmic reticulum (ER), binds peptides to human leukocyte antigen (HLA) class I, and transports the peptide-MHC complex through the Golgi complexes to the cell surface, where it is presented to CD8 + T cells. Therefore, MHC I-peptide complexes are critical for infection surveillance by T cells Gruhler and Fruh (2000); Abbas et al. (2014).

When CTL peptides were identified using NetMHCI and NetCTL 1.2 server, a total of 220 CTL (9-mer) ligands were predicted for the four polyproteins from ancestral viral sequences based on their high combinatorial scores. In addition, a total of 416 epitopes of SARS-CoV-2 variants were predicted based on their high score and submetabolized for other parameters. In particular, for the epitopes with the highest number

Table 1
List of predicted CTL epitopes of the SARS-CoV-2 polyprotein S with all variants Alpha, Beta, Gamma, Delta, and Omicron, each peptide sequence and their number of alleles, antigenicity prediction score, allergenicity, immunogenicity score, toxicity and conservancy.

Protein S	Sequence	NetMHCI Alleles	NetCTL Supertypes	VaxiJen Antigenicity	Allertop Allergenicity	IEDB Immunogenicity	ToxinPred Toxicity	IEDB Conservation
Original Strain								
	857 - GLTVLPPLL	1	1	0,6621	NON-ALLERGEN	0,01706	Non-Toxin	96.30 %(339/352)
	413 - GQTGKIADY	1	2	1,4019	NON-ALLERGEN	0,00796	Non-Toxin	94.03 %(331/352)
	699 - LGAENSVAY	1	1	0,4173	NON-ALLERGEN	0,00912	Non-Toxin	97.44 %(343/352)
	628 - QLTPTWRVY	1	2	1,2119	NON-ALLERGEN	0,31555	Non-Toxin	98.30 %(346/352)
	680 - SPRRARSVA	1	2	0,7729	NON-ALLERGEN	0,04020	Non-Toxin	96.30 %(339/352)
	512 - VLSFELLHA	1	1	1,0776	NON-ALLERGEN	0,16070	Non-Toxin	97.73 %(344/352)
	886 - WTFGAGAAL	1	2	0,4918	NON-ALLERGEN	0,19798	Non-Toxin	99.43 %(350/352)
	88 - GVYFASTEK	4	1	0,7112	NON-ALLERGEN	0,09023	Non-Toxin	95.30 % (345/362)
	504 - YQPYRVVVL	2	5	0,5964	NON-ALLERGEN	0,1409	Non-Toxin	98.89 % (358/362)
	68 - HVSGTNGTK	4	1	1,0956	NON-ALLERGEN	0,06339	Non-Toxin	96.13 % (348/362)
	326 - VRFPNITNL	7	2	1,1141	NON-ALLERGEN	0,1748	Non-Toxin	95.52 % (347/362)
	61 - VTWFHAIHV	4	1	0,5426	NON-ALLERGEN	0,38925	Non-Toxin	95.58 % (346/362)
Original Strain/Alpha/Beta/Delta/Gamma/Omicron								
	341 - FNATRFASV	7	1	0,5609	NON-ALLERGEN	0,14872	Non-Toxin	98.35 % (357/363)
	717 - FTISVTTEI	4	3	0,8535	NON-ALLERGEN	0,04473	Non-Toxin	97.80 % (355/363)
	891 - AALQIPFAM	4	2	0,7747	NON-ALLERGEN	0,12066	Non-Toxin	98.90 % (359/363)
	711 - IAIPTNFTI	3	3	0,7052	NON-ALLERGEN	0,18523	Non-Toxin	97.52 % (354/363)
	257 - WTAGAAAYY	14	4	0,6306	NON-ALLERGEN	0,15259	Non-Toxin	91.18 % (331/363)
	713 - IPTNFTISV	4	1	0,8820	NON-ALLERGEN	0,17229	Non-Toxin	97.52 % (354/363)
	506 - PYRVVVLSF	5	1	1,0281	NON-ALLERGEN	0,03138	Non-Toxin	98.90 % (359/363)
	826 - TLADAGFIK	4	1	0,5781	NON-ALLERGEN	0,28158	Non-Toxin	98.90 % (359/363)
	83 - LPFNDGVYF	1	1	0,5593	NON-ALLERGEN	0,11767	Non-Toxin	96.42 % (350/363)
Alpha/Beta/Delta/Gamma/Omicron								
	1015 - AEIRASANL	4	1	0,7082	NON-ALLERGEN	0,00689	Non-Toxin	100 % (10/10)
	611 - YQGVNCTEV	5	1	1,3957	NON-ALLERGEN	0,08675	Non-Toxin	100 % (10/10)
	1100 - HWFVTQRNF	4	1	0746	NON-ALLERGEN	0,0482	Non-Toxin	100 % (10/10)
	1 - FVFLVLLPL	4	4	0,8601	NON-ALLERGEN	0,04076	Non-Toxin	100 % (10/10)
	1224 - IAIVMVTIM	3	1	1,1339	NON-ALLERGEN	0,06312	Non-Toxin	100 % (10/10)
	665 - IGAGICASY	1	2	0,6368	NON-ALLERGEN	0,06201	Non-Toxin	100 % (10/10)
	1208 - YIKWPWYIW	3	1	0,9673	NON-ALLERGEN	0,42524	Non-Toxin	100 % (10/10)
	881 - ITSGWTFGA	2	1	0,4577	NON-ALLERGEN	0,35124	Non-Toxin	100 % (10/10)
	82 - VLPFNDGVY	1	2	0,4642	NON-ALLERGEN	0,1815	Non-Toxin	100 % (10/10)
	1059 - VVFLHVTYV	8	1	1,5122	NON-ALLERGEN	0,1278	Non-Toxin	100 % (10/10)
Alpha/Beta/Gamma								
	407 - RQIAPGQTG	2	1	1,7890	NON-ALLERGEN	0,02859	Non-Toxin	100 % (6/6)
	442 - SKVGGNYNY	1	1	0,9111	NON-ALLERGEN	0,06751	Non-Toxin	100 % (6/6)
	203 - YSKHTPINL	3	1	1,0547	NON-ALLERGEN	0,09845	Non-Toxin	100 % (6/6)
	1213 - WYIWLGFIA	1	1	1,0356	NON-ALLERGEN	0,46375	Non-Toxin	100 % (6/6)
	1180 - KEIDRLNEV	2	1	0,5300	NON-ALLERGEN	0,15852	Non-Toxin	100 % (6/6)
	629 - TPTWRVYST	3	1	0,4605	NON-ALLERGEN	0,22497	Non-Toxin	100 % (6/6)
	763 - NRALTGIAV	3	1	0,5302	NON-ALLERGEN	0,20642	Non-Toxin	100 % (6/6)
	1058 - GVVFLHVTY	3	1	1,4104	NON-ALLERGEN	0,20837	Non-Toxin	100 % (6/6)
	260 - GAAAYYVGY	1	3	0,6604	NON-ALLERGEN	0,09963	Non-Toxin	100 % (6/6)
	261 - AAAYYVGYL	2	1	0,4605	NON-ALLERGEN	0,07068	Non-Toxin	100 % (6/6)

of alleles ($S^{257-265}$, $S^{1059-1067}$, $S^{326-334}$, $S^{506-314}$, and $S^{611-629}$), the binding of 14, 8, 7, 5, and 5 alleles, respectively, was predicted. Epitopes E^{53-67} and E^{54-68} were predicted to bind the highest number of alleles, 12 and 10, respectively. Similarly, M^{72-86} and M^{71-85} were predicted to bind 7 and 6 alleles of the HLA class, II, respectively. Finally, N^{49-64}, $N^{310-324}$, and $N^{311-325}$ were predicted to bind 12, 12, and 11 alleles, respectively. The other epitopes are characterized by an affinity strength between 1 and 4 alleles Table 1.

Speed and efficiency are the advantages of *in silico* screening of genomic information for vaccine development in the post-genomic era Doytchinova and Flower (2007a). However, this approach is dependent on the accuracy of antigen prediction. Thus, the antigenicity of a sequence may be encoded in a subtle and arcane manner that cannot be identified by direct sequence alignment Doytchinova and Flower (2007b).

Epitope analysis with Vaxijen found only 118 epitopes of proteins from precursor viruses to be antigenic. For sequences of S protein variants, 116 epitopes with high antigenicity scores were selected. Of these, only 9 epitopes were common to all variants, representing one of the lowest concordance rates ever reported Abdelmageed et al. (2020); ul Qamar et al. (2020). This suggests that mutations occurred in regions of high antigenic potential between the different strains. The highest values of antigenicity of the amino acid sequences of protein S ($S^{407-415}$, $S^{1059-1067}$, $S^{1058-1066}$, $S^{611-619}$, $S^{1224-1232}$, $S^{326-334}$, S^{68-76}, $S^{203-211}$,

Table 2
List of predicted CTL epitopes of the SARS-CoV-2 polyprotein E with each peptide sequence and their number ofalleles, antigenicity prediction score, allergenicity, immunogenicity score, toxicity and conservancy.

Protein	Sequence	NetMHCI Alleles	NetCTL Supertypes	VaxiJen Antigenicity	Allertop Allergenicity	IEDB Immunogenicity	ToxinPred Toxicity	IEDB Conservation
E	30 - TLAILTALR	2	1	0,7223	NON-ALLERGEN	0.19890	Non-Toxin	87.50 % (14/16)
	29 - VTLAILTAL	2	1	0,6140	NON-ALLERGEN	0.21055	Non-Toxin	87.50 % (14/16)
	20 - FLAFVVFLL	1	4	0,5308	NON-ALLERGEN	0.30188	Non-Toxin	87.50 % (14/16)
	31 - LAILTALRL	1	1	0,8872	NON-ALLERGEN	0.12755	Non-Toxin	87.50 % (14/16)

Table 3
List of predicted CTL epitopes of the SARS-CoV-2 polyprotein M with each peptide sequence and their number ofalleles, antigenicity prediction score, allergenicity, immunogenicity score, toxicity and conservancy.

Protein	Sequence	NetMHCI Alleles	NetCTL Supertypes	VaxiJen Antigenicity	Allertop Allergenicity	IEDB Immunogenicity	ToxinPred Toxicity	IEDB Conservation
M	57 - LWPVTLACF	2	1	1,1590	NON-ALLERGEN	0,06682	Non-Toxin	98,21 % (55/56)
	136 - SELVIGAVI	1	1	0,6409	NON-ALLERGEN	0,25658	Non-Toxin	96,43 % (54/56)
	188 - AGDSGFAAY	1	1	0,9095	NON-ALLERGEN	0,03981	Non-Toxin	96,43 % (54/56)
	138 - LVIGAVILR	3	1	1,1027	NON-ALLERGEN	0,2601	Non-Toxin	94,64 % (53/56)
	23 - VIGFLFLTW	1	1	1,1465	NON-ALLERGEN	0,24152	Non-Toxin	92,86 % (52/56)
	65 - FVLAAVYRI	2	1	0,5136	NON-ALLERGEN	0,13985	Non-Toxin	92,86 % (52/56)
	94 - SYFIASFRL	2	2	0,4821	NON-ALLERGEN	0,18333	Non-Toxin	91,07 % (51/56)
	50 - KLIFLWLLW	1	1	0,4968	NON-ALLERGEN	0,34287	Non-Toxin	89,29 % (50/56)
	46 - LYIIKLIFL	2	1	0,4865	NON-ALLERGEN	0,1374	Non-Toxin	87,50 % (49/56)
	44 - RFLYIIKLI	1	1	0,4257	NON-ALLERGEN	0,05908	Non-Toxin	87,50 % (49/56)

Table 4
List of predicted CTL epitopes of the SARS-CoV-2 polyprotein N with each peptide sequence and their number ofalleles, antigenicity prediction score, allergenicity, immunogenicity score, toxicity and conservancy.

Protein	Sequence	NetMHCI Alleles	NetCTL Supertypes	VaxiJen Antigenicity	Allertop Allergenicity	IEDB Immunogenicity	ToxinPred Toxicity	IEDB Conservation
N	104 - LSPRWYFYY	1	5	1,2832	NON-ALLERGEN	0,35734	Non-Toxin	97.92 % (141/144)
	105 - SPRWYFYYL	2	2	0,7340	NON-ALLERGEN	0,34101	Non-Toxin	97.92 % (141/144)
	103 - DLSPRWYFY	2	3	1,7645	NON-ALLERGEN	0,25933	Non-Toxin	97.92 % (141/144)
	316 - GMSRIGMEV	2	1	0,6287	NON-ALLERGEN	0,07018	Non-Toxin	97.22 % (140/144)
	323 - EVTPSGTWL	2	1	0,4548	NON-ALLERGEN	0,03442	Non-Toxin	97.22 % (140/144)
	78 - SSPDDQIGY	2	3	0,5260	NON-ALLERGEN	0,0634	Non-Toxin	95.14 % (137/144)
	79 - SPDDQIGYY	3	1	0,4863	NON-ALLERGEN	0,06844	Non-Toxin	94.44 % (136/144)
	361 - KTFPPTEPK	5	1	0,7571	NON-ALLERGEN	0,1306	Non-Toxin	93.06 % (134/144)

$S^{1213-1221}$, $S^{506-514}$, $S^{1208-116}$, $S^{442-450}$, $S^{713-721}$, S^{1-9}, $S^{717-725}$, $S^{891-899}$, $S^{1100-1108}$, S^{88-96}, $S^{1015-1023}$, $S^{711-719}$, $S^{260-268}$, $S^{665-673}$, $S^{257-265}$, $S^{504-512}$, $S^{826-834}$, $S^{341-349}$, S^{83-91}, S^{61-69}, $S^{763-771}$, $S^{1180-1188}$, S^{82-90}, $S^{629-637}$, $S^{261-269}$ and $S^{881-889}$) were 1.7890, 1.5122, 1.4104, 1.3957, 1.1339, 1.1141, 1.0956, 1.0547, 1.0356, 1.0281, 0,9673, 0.9111, 0.8820, 0,8601, 0,8535, 0,7747, 0.7460, 0.7112, 0.7082, 0.7052, 0.6604, 0.6368, 0.6306, 0.5964, 0.5781, 0.5609, 0.5593, 0.5426, 0.5302, 0.5300, 0.4642, 0.4605, 0.4605, and 0.4577 (Table 1).

For the protein E (M), the scores of E^{61-69}, E^{31-39}, E^{55-63}, E^{49-57}, E^{30-38}, E^{57-65}, E^{29-37}, E^{20-28}, E^{51-59} and E^{50-58} (M^{57-65}, M^{23-31}, $M^{138-146}$, M^{84-92}, $M^{188-196}$, $M^{172-180}$, $M^{171-179}$, $M^{170-178}$, M^{37-45}, $M^{136-144}$, $M^{108-116}$, M^{65-73}, M^{50-58}, M^{46-54}, M^{94-102}, M^{54-62} and M^{44-52}) were 0.8998, 0.8872, 0.8251, 0.7476, 0.7223, 0.7020, 0.6140, 0.5308, 0.4213 and 0.4140 (1.1590, 1.1465, 1.1027, 0.9457, 0.9095, 0.7889, 0.7785, 0.7633, 0.7197, 0.6409, 0.6108, 0.5136, 0.4968, 0.4865, 0.4821, 0.4811 and 0.4257), respectively - see details in Table 2 and Table 3.

Finally the amino acids sequences of protein N N^{32-41}, $N^{103-111}$, $N^{100-108}$, $N^{104-112}$, $N^{183-191}$, $N^{345-353}$, $N^{187-195}$, $N^{386-394}$, N^{53-61}, $N^{249-257}$, N^{66-74}, $N^{361-369}$, $N^{322-330}$, $N^{305-313}$, $N^{379-387}$, $N^{105-113}$, $N^{240-248}$, $N^{316-324}$, $N^{266-274}N^{306-314}$, N^{78-86}, N^{79-87} and $N^{323-331}$ were 1.7874, 1.7645, 1.7462, 1.2832, 1.2286, 1.1677, 1.1218, 0.9285, 0.8510, 0.7679, 0.7585, 0.7571, 0.7550, 0.7468, 0.7432, 0.7340, 0.6709, 0.6287, 0.5669, 0.5495, 0.5260, 0.4863 and 0.4548, respectivel (Table 4).

The final peptides were selected and submitted for further analysis. Because of the high current incidence, allergenicity was checked to ensure that the vaccine candidate would not induce allergic reactions once introduced into uniform vectors of equal length using an alignment-independent protein presentation method based on the key physiochemical properties of protein sequences Dimitrov et al. (2013). Of 118 antigenic epitopes, 67 were found to be nonallergenic (Tables 1, 2,3, and 4).

One of the main concerns in subunit vaccine development is the specificity of the immunogenic epitopes. In case, the immunogenicity step is important to determine whether a peptide-MHC complex (pMHC) can be an epitope Calis et al. (2013). Here, only 56 nonallergenic epitopes were selected based on high immunogenicity values. The highest immunogenicity values for the S, E, M, and N proteins of SARS-CoV-2 were 0.46375, 0.30188, 0.34287, and 0.34101, respectively. Subsequently, these epitopes were also found to be non-toxic. Thus, the probability of our epitope-based vaccines causing immunogenic and toxic reactions is very low.

A look at the S protein sequences shows that 21 epitopes with high immunogenicity values were present only in the original strain of SARS-CoV-2, 9 were common to all subvariants and 19 in the Omicron variant, namely $S^{341-349}$, $S^{717-725}$, $S^{891-899}$, $S^{711-719}$, $S^{257-765}$, $S^{713-721}$, $S^{506-514}$, $S^{826-834}$, S^{83-91}, $S^{1015-1023}$, $S^{611-619}$, $S^{1100-1108}$, S^{1-8}, $S^{1224-1232}$, $S^{665-673}$, $S^{1208-1216}$, $S^{881-889}$, S^{82-90} and $S^{1059-1067}$.

Of the 63 epitopes with the best antigenicity, allergenicity, immunogenicity, and toxicity values, 56 obtained conservation index above 90 %. These results suggest that most epitopes in SARS-CoV-2 are highly conserved and are good targets for vaccine development, as these epitopes are regions that evolve slowly and can be expected to be present independently of a particular pathogen strain Bui et al. (2007).

In order to predict whether virus-specific T cells recognize variants, we determined the epitopes with had strong binding affinity with MHC-I. Those with the highest conservation rates more conserved in all variants would be in the construction of the vaccine.

3.3. Helper T Lymphocyte (HTL)

The MHC Class I and Class II epitopes can be recognized by T cell, due to its antigenic nature and acknowledged by the T cell receptors (TCR). MHC class II molecules contain exogenous antigens or pathogen's surface proteins, processed through endocytic pathways to assist the T lymphocytes or CD4 +T cells Reynisson et al. (2020). Thus, with powerful tools for analysis of prediction of T cell epitopes a total of 931 HTL epitopes were predicted from the four SARS-CoV-2 virus proteins using NetMHCII and NetMHCIIpan 4.0 server. In addition, a total of 558 epitopes of all variants with the highest number of alleles were predicted

Table 5
List of predicted HTL epitopes of the SARS-CoV-2 polyprotein S with all variants Alpha, Beta, Gamma, Delta, and Omicron, each peptide sequence and their number of alleles, antigenicity prediction score, allergenicity, Epitope IFN-γ, toxicity and conservancy.

Protein S	Peptide sequence	NetMHCII/NetMHCIIpan Alleles	VaxiJen Antigenicity	Allertop Allergenicity	IFNepitope Epitope IFN-γ	ToxinPred Toxicity	IEDB Conservation
Original Strain							
	139 - PFLGVYYHKNNKSWM	2	0,6641	NON-ALLERGEN	0.36453817	Non-Toxin	89.23 % (315/353)
	142 - GVYYHKNNKSWMESE	2	0,4684	NON-ALLERGEN	0.36198865	Non-Toxin	89.23 % (315/353)
	141 - LGVYYHKNNKSWMES	2	0,4937	NON-ALLERGEN	0.40530432	Non-Toxin	90.08 % (318/353
	257 - GWTAGAAAYYVGYLQ	4	0,5669	NON-ALLERGEN	0.71724775	Non-Toxin	90.93 % (321/353)
	168 - FEYVSQPFLMDLEGK	2	0,8278	NON-ALLERGEN	0.44798419	Non-Toxin	93.76 % (331/353)
Original Strain/Alpha/Beta/Delta/Gamma							
	763 - LNRALTGIAVEQDKN	2	0,4710	NON-ALLERGEN	0,58194087	Non-Toxin	98.61 % (357/362)
	761 - TQLNRALTGIAVEQD	2	0,4153	NON-ALLERGEN	0,43169637	Non-Toxin	98.61 % (357/362)
	59 - FSNVTWFHAIHVSGT	3	0,7533	NON-ALLERGEN	0,094599621	Non-Toxin	94.47 % (342/362)
	431 - GCVIAWNSNNLDSKV	3	0,4585	NON-ALLERGEN	0,022799683	Non-Toxin	95.58 % (346/362)
Original Strain/Alpha/Beta/Delta/Gamma/Omicron							
	895 - QIPFAMQMAYRFNGI	3	0,9573	NON-ALLERGEN	-0,17853133	Non-Toxin	98.62 % (358/363)
	894 - LQIPFAMQMAYRFNG	2	0,7205	NON-ALLERGEN	-0,21947589	Non-Toxin	98.62 % (358/363)
	885 - GWTFGAGAALQIPFA	5	0,4665	NON-ALLERGEN	-0,24676874	Non-Toxin	98.62 % (358/363)
	797 - FGGFNFSQILPDPSK	2	0,4404	NON-ALLERGEN	-0,4760424	Non-Toxin	96.42 % (350/363)
	716 - TNFTISVTTEILPVS	4	1,1691	NON-ALLERGEN	-0,1175069	Non-Toxin	97.25 % (353/363)
	715 - PTNFTISVTTEILPV	5	1,1349	NON-ALLERGEN	0,006286322	Non-Toxin	97.25 % (353/363)
	713 - AIPTNFTISVTTEIL	3	0,6806	NON-ALLERGEN	-0,20095993	Non-Toxin	97.52 % (354/363)
	712 - IAIPTNFTISVTTEI	3	0,7719	NON-ALLERGEN	-0,019108711	Non-Toxin	97.52 % (354/363)
	663- DIPIGAGICASYQTQ	2	1,1088	NON-ALLERGEN	-0,12625968	Non-Toxin	96.42 % (350/363)
	634 - RVYSTGSNVFQTRAG	4	0,4544	NON-ALLERGEN	0,10536516	Non-Toxin	97.25 % (353/363)
	632 - TWRVYSTGSNVFQTR	3	0,4253	NON-ALLERGEN	0,10894371	Non-Toxin	97.52 % (354/363)
	346 - RFASVYAWNRKRISN	5	0,4243	NON-ALLERGEN	0,61013335	Non-Toxin	97.80 % (355/363)
	324 - ESIVRFPNITNLCPF	2	0,6125	NON-ALLERGEN	-0,43992195	Non-Toxin	95.32 % (346/363)
	255 - SSGWTAGAAAYYVGY	2	0,4136	NON-ALLERGEN	0,87705602	Non-Toxin	89.53 % (325/363)
	233 - INITRFQTLLALHRS	2	0,4118	NON-ALLERGEN	-0,007153463	Non-Toxin	91.46 % (332/363)
	232 - GINITRFQTLLALHR	6	0,5582	NON-ALLERGEN	0,10772613	Non-Toxin	91.18 % (331/363)
	231 - IGINITRFQTLLALH	5	0,8391	NON-ALLERGEN	-0,38029989	Non-Toxin	91.46 % (332/363)
	230 - PIGINITRFQTLLAL	2	0,8877	NON-ALLERGEN	-0,042411043	Non-Toxin	91.18 % (331/363)
	166 - CTFEYVSQPFLMDLE	4	0.5700	NON-ALLERGEN	0,043175621	Non-Toxin	92.01 % (334/363)
	165 - NCTFEYVSQPFLMDL	2	0,5206	NON-ALLERGEN	-0,35483268	Non-Toxin	91.74 % (333/363)
Original Strain/Alpha/Beta/Omicron							
	718 - FTISVTTEILPVSMT	3	1,2603	NON-ALLERGEN	-0.40760154	Non-Toxin	99.15 % (353/356)
	1210 - IKWPWYIWLGFIAGL	4	0,9153	NON-ALLERGEN	0.61171913	Non-Toxin	99.43 % (354/356)
	509 - RVVVLSFELLHAPAT	6	0,7485	NON-ALLERGEN	0.5092653	Non-Toxin	98.87 % (352/356)
	2 - FVFLVLLPLVSSQCV	6	0,7185	NON-ALLERGEN	0.092039768	Non-Toxin	97.47 % (347/356)
	238 - FQTLLALHRSYLTPG	2	0,5789	NON-ALLERGEN	0.26071055	Non-Toxin	93.53 % (333/356)
	1 - MFVFLVLLPLVSSQC	7	0,5741	NON-ALLERGEN	0.084372674	Non-Toxin	97.47 % (347/356)
	237 - RFQTLLALHRSYLTP	4	0,5470	NON-ALLERGEN	-0.022265746	Non-Toxin	94.10 % (335/356)
	1103 - FVTQRNFYEPQIITT	2	0,5314	NON-ALLERGEN	-0.4029137	Non-Toxin	100.00 % (356/356)
	52 - QDLFLPFFSNVTWFH	3	0,4159	NON-ALLERGEN	-0.051186129	Non-Toxin	95,78 % (341/356)
	1060 - VVFLHVTYVPAQEKN	2	1,1720	NON-ALLERGEN	0.035308408	Non-Toxin	100.00 % (356/356)
	1059 - GVVFLHVTYVPAQEK	5	1,1043	NON-ALLERGEN	0.25556948	Non-Toxin	100.00 % (356/356)
Alpha/Beta/Gamma/Omicron							
	1061 - VFLHVTYVPAQEKNF	4	1,0339	NON-ALLERGEN	-0.070782391	Non-Toxin	100.00 % (7/7)
	511 - VVLSFELLHAPATVC	5	0,8618	NON-ALLERGEN	0.47729732	Non-Toxin	100.00 % (7/7)
	508 - YRVVVLSFELLHAPA	2	0,7072	NON-ALLERGEN	0.77626818	Non-Toxin	100.00 % (7/7)
	886 - WTFGAGAALQIPFAM	2	0,6670	NON-ALLERGEN	-0.38373216	Non-Toxin	100.00 % (7/7)
	1214 - WYIWLGFIAGLIAIV	4	0,5770	NON-ALLERGEN	10.656.151	Non-Toxin	100.00 % (7/7)
	512 - VLSFELLHAPATVCG	2	0,4784	NON-ALLERGEN	0.096827098	Non-Toxin	100.00 % (7/7)
Alpha/Beta/Gamma							
	259 - TAGAAAYYVGYLQPR	2	1,0413	NON-ALLERGEN	0.55255681	Non-Toxin	100.00 % (4/4)
	754 - LQYGSFCTQLNRALT	2	1027	NON-ALLERGEN	0.1500564	Non-Toxin	100.00 % (4/4)
	628 - QLTPTWRVYSTGSNV	2	0,9276	NON-ALLERGEN	-0.24711649	Non-Toxin	100.00 % (4/4)
	260 - AGAAAYYVGYLQPRT	2	0,9134	NON-ALLERGEN	0.59506751	Non-Toxin	100.00 % (4/4)
	1218 - LGFIAGLIAIVMVTI	2	0,8933	NON-ALLERGEN	0.32119174	Non-Toxin	100.00 % (4/4)
	751 - NLLLQYGSFCTQLNR	2	0,8668	NON-ALLERGEN	0.1651956	Non-Toxin	100.00 % (4/4)
	750 - SNLLLQYGSFCTQLN	2	0,8305	NON-ALLERGEN	0.27039446	Non-Toxin	100.00 % (4/4)
	890 - AGAALQIPFAMQMAY	2	0,8216	NON-ALLERGEN	-0.40373248	Non-Toxin	100.00 % (4/4)
	325 - SIVRFPNITNLCPFG	2	0,7899	NON-ALLERGEN	-0.23626945	Non-Toxin	100.00 % (4/4)
	312 - IYQTSNFRVQPTESI	2	0,7459	NON-ALLERGEN	0.30406292	Non-Toxin	100.00 % (4/4)
	760 - CTQLNRALTGIAVEQ	3	0,7454	NON-ALLERGEN	0.16758267	Non-Toxin	100.00 % (4/4)
	1212 - WPWYIWLGFIAGLIA	2	0,7293	NON-ALLERGEN	11.863.034	Non-Toxin	100.00 % (4/4)
	591 - SFGGVSVITPGTNTS	2	0,6553	NON-ALLERGEN	-0.21404582	Non-Toxin	100.00 % (4/4)
	218 - QGFSALEPLVDLPIG	2	0,6177	NON-ALLERGEN	-0.044834443	Non-Toxin	100.00 % (4/4)
	1216 - IWLGFIAGLIAIVMV	6	0615	NON-ALLERGEN	0.86284921	Non-Toxin	100.00 % (4/4)
	1215 - YIWLGFIAGLIAIVM	5	0609	NON-ALLERGEN	0.89392418	Non-Toxin	100.00 % (4/4)
	50 - STQDLFLPFFSNVTW	2	0,6034	NON-ALLERGEN	0.043916948	Non-Toxin	100.00 % (4/4)
	1219 - GFIAGLIAIVMVTIM	2	0,5098	NON-ALLERGEN	0.077612198	Non-Toxin	100.00 % (4/4)
	55 - FLPFFSNVTWFHAIH	2	0,4883	NON-ALLERGEN	0.17628387	Non-Toxin	100.00 % (4/4)
	140 - FLGVYYHKNNKSWME	4	0,4793	NON-ALLERGEN	0.53199532	Non-Toxin	100.00 % (4/4)

(*continued on next page*)

Table 5 (*continued*)

Protein S	Peptide sequence	NetMHCII/NetMHCIIpan Alleles	VaxiJen Antigenicity	Allertop Allergenicity	IFNepitope Epitope IFN-γ	ToxinPred Toxicity	IEDB Conservation
	368 - LYNSASFSTFKCYGV	2	0,4171	NON-ALLERGEN	0.25053962	Non-Toxin	100.00 % (4/4)
	1259 - DDSEPVLKGVKLHYT	2	1,1849	NON-ALLERGEN	-0.70684166	Non-Toxin	100.00 % (4/4)
	892 - AALQIPFAMQMAYRF	5	0,9108	NON-ALLERGEN	-0.71321984	Non-Toxin	100.00 % (4/4)
	749 - CSNLLLQYGSFCTQL	3	0,6336	NON-ALLERGEN	0.2869462	Non-Toxin	100.00 % (4/4)
	264 - AYYVGYLQPRTFLLK	6	0,4269	NON-ALLERGEN	0.31280017	Non-Toxin	100.00 % (4/4)

Table 6
List of predicted HTL epitopes of the SARS-CoV-2 polyprotein M each peptide sequence and their number of alleles, antigenicity prediction score, allergenicity, Epitope IFN-γ, toxicity and conservancy.

Protein	Sequence	NetMHCII/NetMHCIIpan Alleles	VaxiJen Antigenicity	Allertop Allergenicity	IFNepitope Epitope IFN-γ	ToxinPred Toxicity	IEDB Conservation
M	136 - SELVIGAVILRGHLR	2	0,6768	NON-ALLERGEN	0.67529222	Non-Toxin	94.64 % (53/56)
	72 - RINWITGGIAIAMAC	7	1,1629	NON-ALLERGEN	0.097493902	Non-Toxin	89.29 % (50/56)
	71 - YRINWITGGIAIAMA	6	1,1274	NON-ALLERGEN	-0.032197757	Non-Toxin	89.29 % (50/56)
	34 - LLQFAYANRNRFLYI	4	0,7387	NON-ALLERGEN	-0.59982611	Non-Toxin	85.71 % (48/56)

Table 7
List of predicted HTL epitopes of the SARS-CoV-2 polyprotein E each peptide sequence and their number of alleles, antigenicity prediction score, allergenicity, Epitope IFN-γ, toxicity and conservancy.

Protein	Sequence	NetMHCII/NetMHCIIpan Alleles	VaxiJen Antigenicity	Allertop Allergenicity	IFNepitope Epitope IFN-γ	ToxinPred Toxicity	IEDB Conservation
E	53 - KPSFYVYSRVKNLNS	12	0,8229	NON-ALLERGEN	-0.52339177	Non-Toxin	68.75 % (11/16)
	52 - VKPSFYVYSRVKNLN	9	1,2319	NON-ALLERGEN	-0.46468316	Non-Toxin	68.75 % (11/16)
	51 - LVKPSFYVYSRVKNL	6	0,7311	NON-ALLERGEN	-0.39721306	Non-Toxin	68.75 % (11/16)
	56 - FYVYSRVKNLNSSRV	5	0,6103	NON-ALLERGEN	-0.1032629	Non-Toxin	62.50 % (10/16)
	54 - PSFYVYSRVKNLNSS	10	0,7986	NON-ALLERGEN	-0.35688916	Non-Toxin	56.25 % (9/16)
	55 - SFYVYSRVKNLNSSR	8	0,6291	NON-ALLERGEN	-0.4195908	Non-Toxin	56.25 % (9/16)
	57 - YVYSRVKNLNSSRVP	3	0,4492	NON-ALLERGEN	-0.22463037	Non-Toxin	50.00 % (8/16)

Table 8
List of predicted HTL epitopes of the SARS-CoV-2 polyprotein N each peptide sequence and their number of alleles, antigenicity prediction score, allergenicity, Epitope IFN-γ, toxicity and conservancy.

Protein	Peptide sequence	NetMHCII/NetMHCIIpan Alleles	VaxiJen Antigenicity	Allertop Allergenicity	IFNepitope Epitope IFN-γ	ToxinPred Toxicity	IEDB Conservation
N	108 - PSASAFFGMSRIGME	4	0,6408	NON-ALLERGEN	0,94891009	Non-Toxin	97.22 % (140/144)
	83 - QIAQFAPSASAFFGM	6	0,4032	NON-ALLERGEN	0,83655244	Non-Toxin	98.61 % (142/144)
	259 - SASAFFGMSRIGMEV	3	0,6584	NON-ALLERGEN	0,79707565	Non-Toxin	97.22 % (140/144)
	310 - QIGYYRRATRRIRGG	12	0,4614	NON-ALLERGEN	0,70344772	Non-Toxin	95.83 % (138/144)
	311 - IGYYRRATRRIRGGD	11	0,6649	NON-ALLERGEN	0,6416295	Non-Toxin	95.83 % (138/144)
	386 - QRQKKQQTVTLLPAA	1	0,6824	NON-ALLERGEN	0,13245676	Non-Toxin	95.14 % (137/144)
	264 - ASAFFGMSRIGMEVT	3	0,8620	NON-ALLERGEN	0,085634049	Non-Toxin	97.22 % (140/144)
	330 - QTVTLLPAADLDDFS	4	0,5213	NON-ALLERGEN	0,072849059	Non-Toxin	95.83 % (138/144)
	49 - TASWFTALTQHGKED	12	0,4491	NON-ALLERGEN	-0,079044179	Non-Toxin	98.61 % (142/144)
	84 - RQKRTATKAYNVTQA	3	0,6318	NON-ALLERGEN	-0,095407182	Non-Toxin	97.92 % (141/144)
	389 - RQKKQQTVTLLPAAD	2	0,6555	NON-ALLERGEN	-0,098541454	Non-Toxin	94.44 % (136/144)
	266 - GTWLTYTGAIKLDDK	6	0,9934	NON-ALLERGEN	-0,13745422	Non-Toxin	96.53 % (139/144)
	390 - QKKQQTVTLLPAADL	4	0,7662	NON-ALLERGEN	-0,14422474	Non-Toxin	94.44 % (136/144)
	303 - TWLTYTGAIKLDDKD	2	1,2416	NON-ALLERGEN	-0,21623699	Non-Toxin	96.53 % (139/144)
	329 - QQTVTLLPAADLDDF	4	0,4614	NON-ALLERGEN	-0,24539346	Non-Toxin	95.83 % (138/144)
	384 - RWYFYYLGTGPEAGL	2	0,7505	NON-ALLERGEN	-0,28690274	Non-Toxin	95.14 % (137/144)
	328 - SGTWLTYTGAIKLDD	4	0,6215	NON-ALLERGEN	-0,3048811	Non-Toxin	95.83 % (138/144)
	105 - ATKAYNVTQAFGRRG	4	0,7146	NON-ALLERGEN	-0,41990668	Non-Toxin	97.92 % (141/144)
	327 - PRWYFYYLGTGPEAG	2	0,8083	NON-ALLERGEN	-0,46153318	Non-Toxin	95.83 % (138/144)
	309 - WLTYTGAIKLDDKDP	1	1,2787	NON-ALLERGEN	-0,5483585	Non-Toxin	96.53 % (139/144)
	265 - SPRWYFYYLGTGPEA	1	0,8767	NON-ALLERGEN	-0,61601606	Non-Toxin	96.53 % (139/144)
	107 - KAYNVTQAFGRRGPE	1	0,6104	NON-ALLERGEN	-0,76630992	Non-Toxin	97.92 % (141/144)
	385 - WYFYYLGTGPEAGLP	3	0,7188	NON-ALLERGEN	-0,88957872	Non-Toxin	95.14 % (137/144)
	106 - TKAYNVTQAFGRRGP	5	0,5975	NON-ALLERGEN	-0,89365915	Non-Toxin	97.92 % (141/144)

and submetabolized for other parameters. Specifically, for proteins S and E, the epitopes $S^{1014-1028}$, $S^{795-809}$, $S^{690-704}$, $S^{310-324}$, E^{53-67}, and E^{54-68} were predicted to bind 12, 10, 9, 8 12, and 10 alleles, respectively. Similarly, M^{72-86} and M^{71-85} (N^{49-64}, $N^{310-324}$, and $N^{311-325}$) were predicted to bind 7 and 6 (12, 12, and 11) alleles of the HLA class II (see details in Table 5, Table 6, Table 7, and Table 8).

Because antigenicity is the ability to recognize a specific antigen accompanied by an immune response, antigenic epitopes that can be recognized by host immune cells and elicit both a humoral and a cellular immune response to the viral antigen could be used as potential vaccine

Table 9
Population coverage by MHC-I ($CD8^+$ T cell) epitopes.

	IEDB					IEDB			
Protein	Sequence	North America	South America	Average	Protein	Sequence	North America	South America	Average
	Original Strain					Original Strain/Alpha/Beta/Delta/Gamma			
E	FLAFVVFLL	42.82 %	21.41 %	32.12 %		GVYFASTEK	8.94 %	3.10 %	6.02 %
	TLAILTALR	9.22 %	14.30 %	11.76 %		YQPYRVVVL	11.77 %	8.94 %	10.36 %
	LAILTALRL	8.37 %	3.96 %	6.17 %		HVSGTNGTK	23.66 %	11.42 %	17.54 %
M	VTLAILTAL	5.33 %	1.84 %	3.59 %		VTWFHAIHV	0.19 %	0.00 %	0.10 %
	LVIGAVILR	15.97 %	29.57 %	22.77 %		Original Strain/Alpha/Beta/Delta/Gamma/Omicron			
	LYIIKLIFL	30.06 %	27.61 %	28.84 %		IGAGICASY	5.11 %	3.14 %	4.13 %
	LWPVTLACF	30.06 %	27.61 %	28.84 %		AEIRASANL	26.88 %	18.96 %	22.92 %
	FVLAAVYRI	44.14 %	21.85 %	33.00 %		IAIVMVTIM	20.11 %	5.15 %	12.63 %
	SYFIASFRL	30.06 %	27.61 %	28.84 %		VLPFNDGVY	5.11 %	3.14 %	4.13 %
	VIGFLFLTW	3.63 %	1.35 %	2.49 %		VRFPNITNL	53.93 %	48.29 %	51.11 %
	RFLYIIKLI	8.21 %	4.42 %	6.32 %		VVFLHVTYV	63.33 %	42.17 %	52.75 %
	KLIFLWLLW	3.63 %	1.35 %	2.49 %		ITSGWTFGA	5.13 %	7.74 %	6.44 %
	SELVIGAVI	9.61 %	2.81 %	6.21 %		YIKWPWYIW	13.76 %	8.27 %	11.02 %
N	AGDSGFAAY	12.72 %	6.03 %	9.38 %		YQGVNCTEV	1.74 %	23.97 %	12.86 %
	KTFPPTEPK	17.71 %	7.92 %	12.82 %		AALQIPFAM	13.41 %	3.78 %	8.60 %
	SPDDQIGYY	30.10 %	11.07 %	20.59 %		FNATRFASV	27.37 %	9.21 %	18.29 %
	SSPDDQIGY	17.71 %	7.92 %	12.82 %		FTISVTTEI	6.75 %	12.84 %	9.80 %
	DLSPRWYFY	12.72 %	6.03 %	9.38 %		FVFLVLLPL	42.67 %	46.56 %	44.62 %
	SPRWYFYYL	21.04 %	7.77 %	14.41 %		IAIPTNFTI	6.45 %	2.24 %	4.35 %
	GMSRIGMEV	42.97 %	21.41 %	32.19 %		IPTNFTISV	6.42 %	7.06 %	6.74 %
	EVTPSGTWL	9.95 %	8.47 %	9.21 %		LPFNDGVYF	17.71 %	7.92 %	12.82 %
	LSPRWYFYY	40.44 %	39.43 %	39.94 %		PYRVVVLSF	13.41 %	3.78 %	8.60 %
S	AAAYYVGYL	6.48 %	5.08 %	5.78 %		TLADAGFIK	12.70 %	4.75 %	8.73 %
	TPTWRVYST	4.93 %	2.01 %	3.47 %		HWFVTQRNF	29.33 %	33.71 %	31.52 %
	NRALTGIAV	7.85 %	4.90 %	6.38 %	S	WTAGAAAYY	6.45 %	2.24 %	4.35 %
	KEIDRLNEV	6.85 %	12.30 %	9.58 %					
	SKVGGNYNY	1.65 %	0.00 %	0.83 %					
	WYIWLGFIA	0.07 %	0.00 %	0.04 %					
	YSKHTPINL	17.49 %	14.45 %	15.97 %					
	RQIAPGQTG	7.87 %	6.16 %	7.02 %					
	GVVFLHVTY	6.02 %	5.78 %	5.90 %					
	GAAAYYVGY	5.48 %	4.17 %	4.83 %					

targets Doytchinova and Flower (2007b). To this end, 169 epitopes were found to be antigens when analyzed with Vaxijen. For the sequence of all variants of the S protein, 320 epitopes with a high antigenicity score were selected. In contrast, 95 epitopes were identical to the CTL epitopes of the original strain. The highest antigenicity scores for the amino acid sequences of protein S ($S^{1060-1074}$, $S^{716-730}$, $S^{715-729}$, $S^{663-677}$, $S^{1059-1074}$, and $S^{1061-1075}$) were 1.1720, 1.1691, 1.1349, 1.1088, 1.1043, and 1.0339. For protein E (E^{53-67}, E^{53-67}, and E^{54-68}), the values were 1.2319, 0.8229, and 0.7986, respectively. Similarly, the values for protein M (M^{72-86} and M^{71-85}) were 1.1629 and 1.1274, respectively. In addition, the values for protein N ($N^{309-325}$ and $N^{303-317}$) were 1.2787 and 1.2416, respectively.

These epitopes have been subjected to allergenicity and toxicity analysis because they represent an important hurdle in vaccine development. Therefore, Allertop and Toxinpred aim to predict allergens and non-allergens with high sensitivity and specificity Dimitrov et al. (2013). The servers identified 130 HTL epitopes (15-mer) as likely non-allergens and non-toxic. This type of analysis saves time, resources, and money for the pharmaceutical and vaccine industries Gupta et al. (2013). For the S protein, 160 epitopes were selected as nonallergenic and nontoxic, of which 50 epitopes are present in the wild-type strain. Therefore, the likelihood of our epitope-based vaccines causing allergic and toxic reactions is very low. Finally, we identified the 7 mutations with strong binding affinity, high antigenicity, allergenicity, immunogenicity, and nontoxicity present in protein S of all SARS-CoV-2 variants that we could readily use in the construction of our prototype vaccine.

A total of 75 epitopes of the original strain have a conservation of more than 90 %, but only 20 epitopes are present in all variants. These values suggest that the final epitopes are regions that are good targets for vaccine development because the sequences are likely to be conserved regardless of disease stage or a particular pathogen strain, making them effective vaccines. Bui et al. (2007).

3.4. *IFN-γ inducing epitope prediction*

A total of 106 potential IFN-γ inducing epitopes (15-mer) were predicted by the IFNepitope server. Of these, only 55 proved to be IFN-γ positive epitopes. These epitopes were selected on the basis of the high percentile for further analysis of overlap with B-cell epitopes. This validation is important because HTL epitopes that release cytokines such as interferon-gamma (IFN-γ) have been shown to be an important mediator of protection against SARS-CoV-2 Lagunas-Rangel and Chávez-Valencia (2020). After immunization, there is production of IFN-γ and a consistent increase in the Th (helper) cell population with memory development Seder et al. (2008). Therefore, IFN-/*gamma* plays an important role in the clearance of viral infection Chesler and Reiss (2002).

3.5. *Population coverage*

The frequency of HLA genotype changes in different populations of the world and the nature of HLA polymorphism affects the binding of a peptide identified as an "epitope" during vaccine development, in part because the binding of peptides to MHC molecules is allele-specific Bui et al. (2006). Because we know that selecting multiple peptides of SARS-CoV-2 with different HLA binding specificities will provide greater coverage of the patient population targeted by our SARS-CoV-2 multi-epitope vaccine, the IEDB Population Coverage server was used to calculate the population coverage value of each peptide in different geographic regions based on their MHC binding alleles Bui et al. (2006).

In this study, the MHC-I-binding alleles of 56 epitopes were identified, mainly alleles common in populations from North and South American regions (Table 9). The highest and lowest population coverage in North America were 63.33 % and 0.07 % for $S^{1059-1073}$ and $S^{1213-1227}$, respectively. The highest (lowest) value of population

Table 10
Population coverage by MHC-II ($CD4^+$ T cell) epitopes.

	IEDB					IEDB			
Protein	Sequence	North America	South America	Average	Protein	Sequence	North America	South America	Average
	Original Strain					**Original Strain/Alpha/Beta**			
E	PSFYVYSRVKNLNSS	99.98 %	99.44 %	99.71 %		VVFLHVTYVPAQEKN	35.60 %	57.57 %	46.59 %
	KPSFYVYSRVKNLNS	99.98 %	99.24 %	99.61 %		GVVFLHVTYVPAQEK	67.03 %	59.26 %	63.15 %
	VKPSFYVYSRVKNLN	79.78 %	92.60 %	86.19 %		VFLHVTYVPAQEKNF	99.99 %	99.82 %	99.91 %
	SFYVYSRVKNLNSSR	48.03 %	64.00 %	56.02 %		TAGAAAYYVGYLQPR	99.26 %	68.99 %	84.13 %
	LVKPSFYVYSRVKNL	51.70 %	42.86 %	47.28 %		LQYGSFCTQLNRALT	15.47 %	2.28 %	8.88 %
	FYVYSRVKNLNSSRV	21.11 %	23.71 %	22.41 %		QLTPTWRVYSTGSNV	2.23 %	20.17 %	11.20 %
	YVYSRVKNLNSSRVP	21.46 %	11.15 %	16.31 %		AGAAAYYVGYLQPRT	13.04 %	15.13 %	14.09 %
M	RINWITGGIAIAMAC	97.28 %	98.63 %	97.96 %		LGFIAGLIAIVMVTI	0.0	25.70 %	12.85 %
	YRINWITGGIAIAMA	94.50 %	97.43 %	95.97 %		NLLLQYGSFCTQLNR	1.24 %	1.25 %	1.25 %
	SELVIGAVILRGHLR	84.66 %	77.13 %	80.90 %		SNLLLQYGSFCTQLN	99.26 %	68.99 %	84.13 %
	LLQFAYANRNRFLYI	32.68 %	13.63 %	23.16 %		AGAALQIPFAMQMAY	13.32 %	12.73 %	13.03 %
N	QIGYYRRATRRIRGG	99.98 %	97.25 %	98.66 %		SIVRFPNITNLCPFG	15.47 %	2.28 %	8.88 %
	RQKRTATKAYNVTQA	99.91 %	92.19 %	96.06 %		IYQTSNFRVQPTESI	19.38 %	9.31 %	14.35 %
	IGYYRRATRRIRGGD	99.87 %	92.91 %	96.40 %		CTQLNRALTGIAVEQ	22.34 %	37.89 %	30.12 %
	TASWFTALTQHGKED	98.56 %	99.78 %	99.17 %		WPWYIWLGFIAGLIA	99.26 %	68.99 %	84.13 %
	QKKQQTVTLLPAADL	88.07 %	85.16 %	86.62 %		SFGGVSVITPGTNTS	59.47 %	46.88 %	53.18 %
	QIAQFAPSASAFFGM	85.01 %	87.02 %	86.02 %		QGFSALEPLVDLPIG	13.04 %	15.13 %	14.09 %
	WYFYYLGTGPEAGLP	76.52 %	70.01 %	73.27 %		IWLGFIAGLIAIVMV	79.97 %	60.77 %	70.37 %
	GTWLTYTGAIKLDDK	75.35 %	91.49 %	83.42 %		YIWLGFIAGLIAIVM	99.82 %	90.68 %	95.25 %
	QTVTLLPAADLDDFS	74.96 %	72.42 %	73.69 %		STQDLFLPFFSNVTW	13.04 %	15.13 %	14.09 %
	QQTVTLLPAADLDDF	74.96 %	72.42 %	73.69 %		GFIAGLIAIVMVTIM	5.33 %	7.77 %	6.55 %
	PSASAFFGMSRIGME	67.63 %	77.19 %	72.41 %		FLPFFSNVTWFHAIH	0.0	0.0	0.00 %
	TKAYNVTQAFGRRGP	42.56 %	52.26 %	47.41 %		FLGVYYHKNNKSWME	25.44 %	7.30 %	16.37 %
	SGTWLTYTGAIKLDD	41.80 %	44.97 %	43.39 %		LYNSASFSTFKCYGV	19.38 %	9.31 %	14.35 %
	RQKKQQTVTLLPAAD	36.68 %	28.10 %	32.39 %		FTISVTTEILPVSMT	19.38 %	32.61 %	26.00 %
	ATKAYNVTQAFGRRG	36.54 %	50.20 %	43.37 %		DDSEPVLKGVKLHYT	3.96 %	1.70 %	2.83 %
	RWYFYYLGTGPEAGL	3.96 %	3.16 %	3.56 %		IKWPWYIWLGFIAGL	99.88 %	95.00 %	97.44 %
	PRWYFYYLGTGPEAG	3.96 %	3.16 %	3.56 %		AALQIPFAMQMAYRF	99.45 %	71.11 %	85.28 %
	QRQKKQQTVTLLPAA	27.70 %	19.95 %	23,83 %		VVLSFELLHAPATVC	99.32 %	84.10 %	91.71 %
	TWLTYTGAIKLDDKD	24.15 %	16.70 %	20,43 %		RVVVLSFELLHAPAT	99.91 %	87.89 %	93.90 %
	WLTYTGAIKLDDKDP	19.38 %	9.31 %	14.35 %		FVFLVLLPLVSSQCV	99.97 %	96.08 %	98.03 %
	ASAFFGMSRIGMEVT	16.68 %	7.75 %	12.26 %		YRVVVLSFELLHAPA	13.04 %	15.13 %	14.09 %
	SASAFFGMSRIGMEV	16.68 %	7.75 %	12.26 %		WTFGAGAALQIPFAM	10.17 %	35.26 %	22.72 %
	SPRWYFYYLGTGPEA	1.24 %	1.25 %	1.25 %		CSNLLLQYGSFCTQL	99.43 %	70.60 %	85.02 %
	KAYNVTQAFGRRGPE	0.0 %	0.0 %	0.0 %		FQTLLALHRSYLTPG	19.38 %	9.31 %	14.35 %
S	CDIPIGAGICASYQT	70.98 %	63.08 %	70.98 %		WYIWLGFIAGLIAIV	26.04 %	49.32 %	37.68 %
	FEYVSQPFLMDLEGK	91.41 %	85.45 %	91.41 %		MFVFLVLLPLVSSQC	99.97 %	96.08 %	98.03 %
	FKIYSKHTPINLVRD	99.97 %	98.58 %	99.97 %		RFQTLLALHRSYLTP	23.87 %	30.46 %	27.17 %
	GIVNNTVYDPLQPEL	98.81 %	99.88 %	98.81 %		FVTQRNFYEPQIITT	13.04 %	15.13 %	14.09 %
	GIYQTSNFRVQPTES	16.32 %	40.21 %	16.32 %		VLSFELLHAPATVCG	22.79 %	5.20 %	14.00 %
	GVYYHKNNKSWMESE	75.37 %	66.91 %	75.37 %		AYYVGYLQPRTFLLK	99.90 %	95.21 %	97.56 %
	GWTAGAAAYYVGYLQ	54.63 %	61.17 %	54.63 %		QDLFLPFFSNVTWFH	30.64 %	38.63 %	34.64 %
	HWFVTQRNFYEPQII	99.95 %	96.22 %	99.95 %		**Original Strain/Alpha/Beta/Delta/Gama**			
	IPTNFTISVTTEILP	95.33 %	97.35 %	95.33 %		LNRALTGIAVEQDKN	14.11 %	16.18 %	15.15 %
	ITSGWTFGAGAALQI	49.82 %	54.58 %	49.82 %		TQLNRALTGIAVEQD	29.93 %	33.06 %	31.50 %
	KDFGGFNFSQILPDP	100.00 %	99.53 %	100 %		FSNVTWFHAIHVSGT	29.56 %	24.14 %	26.85 %
	KGIYQTSNFRVQPTE	1.24 %	1.25 %	1.24 %		GCVIAWNSNNLDSKV	99.83 %	88.95 %	94.39 %
	KHTPINLVRDLPQGF	99.99 %	98.93 %	99.99 %		**Original Strain/Alpha/Beta/Delta/Gama/Omicron**			
	LGVYYHKNNKSWMES	8.30 %	3.78 %	8.3 %		QIPFAMQMAYRFNGI	88.37 %	91.57 %	89.97 %
	NGVGYQPYRVVVLSF	31.09 %	10.07 %	31.09 %		LQIPFAMQMAYRFNG	88.37 %	91.57 %	89.97 %
	NRALTGIAVEQDKNT	41.80 %	44.97 %	41.8 %		GWTFGAGAALQIPFA	46.92 %	41.86 %	44.39 %
	NTSNQVAVLYQGVNC	99.96 %	98.03 %	99.96 %		FGGFNFSQILPDPSK	45.94 %	42.30 %	44.12 %
	PFLGVYYHKNNKSWM	99.99 %	99.96 %	99.99 %		TNFTISVTTEILPVS	62.96 %	82.19 %	72.58 %
	QSIIAYTMSLGAENS	49.82 %	54.58 %	49.82 %		PTNFTISVTTEILPV	91.63 %	94.17 %	92.90 %
	RAAEIRASANLAATK	0.0 %	0.0 %	0 %		AIPTNFTISVTTEIL	94.87 %	97.54 %	96.21 %
	TESIVRFPNITNLCP	99.97 %	98.38 %	99.97 %		IAIPTNFTISVTTEI	98.85 %	99.84 %	99.35 %
	TESNKKFLPFQQFGR	99.98 %	99.00 %	99.98 %		DIPIGAGICASYQTQ	70.98 %	63.08 %	67.03 %
	TFEYVSQPFLMDLEG	99.99 %	99.21 %	99.99 %		RVYSTGSNVFQTRAG	99.74 %	99.94 %	99.84 %
	TPINLVRDLPQGFSA	99.97 %	98.58 %	99.97 %		TWRVYSTGSNVFQTR	99.76 %	93.83 %	96.80 %
	TPPIKDFGGFNFSQI	38.89 %	21.78 %	38.89 %		RFASVYAWNRKRISN	99.94 %	98.24 %	99.09 %
	VYSTGSNVFQTRAGC	98.01 %	98.62 %	98.01 %		ESIVRFPNITNLCPF	99.97 %	98.50 %	99.24 %
	WRVYSTGSNVFQTRA	99.64 %	76.09 %	99.64 %		SSGWTAGAAAYYVGY	38.89 %	21.78 %	30.34 %
	YQTSNFRVQPTESIV	8.30 %	3.78 %	8.3 %		INITRFQTLLALHRS	27.70 %	19.95 %	23.83 %
						GINITRFQTLLALHR	33.20 %	11.48 %	22.34 %
						IGINITRFQTLLALH	39.86 %	33.92 %	36.89 %
						PIGINITRFQTLLAL	23.28 %	10.60 %	16.94 %
						CTFEYVSQPFLMDLE	99.99 %	99.18 %	99.59 %
					S	NCTFEYVSQPFLMDL	100.00 %	99.45 %	99.73 %

Table 11
B cells linear epitopes of SARS-CoV-2 structural proteins.

	IEDB				
Protein	Start	End	Peptide	Length	Conservancy
Original Strain					
E	6	9	SEET	4	93.75 %(15/16)
	59	78	YSRVKNLNSSRVPDLLVLPP	20	00.0 %(0/16)
M	5	20	NGTITVEELKKLLEQW	16	89.29 % (50/56)
	40	41	AN	2	98.21 % (55/56)
	132	137	PLLESE	6	96.43 % (54/56)
	161	163	IKD	3	94.64 % (53/56)
	180	191	KLGASQRVAGDS	12	96.43 % (54/56)
	199	218	YRIGNYKLNTDHSSSSDNIA	20	94.64 % (53/56)
N	4	15	NGPQNQRNAPRI	12	95.83 % (138/144)
	17	48	FGGPSDSTGSNQNGERSGARSKQRRPQGLPNN	32	88.89 % (128/144)
	59	105	HGKEDLKFPRGQGVPINTNSSPDDQIGYYRRATRRIRGGDGKMKDLS	47	90.28 % (130/144)
	119	127	AGLPYGANK	9	96.53 % (139/144)
	137	163	GALNTPKDHIGTRNPANNAAIVLQLPQ	27	93.75 % (135/144)
	165	216	TTLPKGFYAEGSRGGSQASSRSSSRSRNSSRNSTPGSSRGTSPARMAGNGGD	52	61.81 % (89/144)
	226	267	RLNQLESKMSGKGQQQQGQTVTKKSAAEASKKPRQKRTATKA	42	88.19 % (127/144)
	276	299	RRGPEQTQGNFGDQELIRQGTDYK	24	97.92 % (141/144)
	343	348	DPNFKD	6	99.31 % (143/144)
	358	402	DAYKTFPPTEPKKDKKKKADETQALPQRQKKQQTVTLLPAADLDD	45	84.03 % (121/144)
	404	416	SKQLQQSMSSADS	13	96.53 % (139/144)
Original Strain/Alpha/Beta/Delta/Gamma/Omicron					
S	13	37	SQCVNLTTRTQLPPAYTNSFTRGVY	25	90.63 % (329/363)
	177	189	MDLEGKQGNFKNL	13	87.33 % (317/363)
	293	296	LDPL	4	95.59 % (347/363)
	329	363	FPNITNLCPFGEVFNATRFASVYAWNRKRISNCVA	35	95.04 % (345/363)
	369	393	YNSASFSTFKCYGVSPTKLNDLCFT	25	95.32 % (346/363)
	555	562	SNKKFLPF	8	99.17 % (360/363)
	602	606	TNTSN	5	99.72 % (362/363)
	617	632	CTEVPVAIHADQLTPT	16	95.87 % (348/363)
	748	748	E	1	100.00 % (363/363)
	773	779	EQDKNTQ	7	98.35 % (357/363)
	828	842	LADAGFIKQYGDCLG	15	98.07 % (356/363)
	1107	1118	RNFYEPQIITTD	12	99.17 % (360/363)
	1133	1172	VNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGI	40	95.87 % (348/363)
	1203	1206	LGKY	4	98.35 % (357/363)

Table 12
Predicted CTL and HTL epitopes among the SARS-CoV-2 virus structural proteins overlapping with B-cell epitopes of the same.

	IEDB		
Protein	CTL epitope	HTL epitope	B-Cell epitope
M		SELVIGAVILRGHLR	PLLESE
N	SSPDDQIGY	QIGYYRRATRRIRGG	
N	SPDDQIGYY	IGYYRRATRRIRGGD	HGKEDLKFPRGQGVPINTNSSPDDQIGYYRRATRRIRGGDGKMKDLS
N	KTFPPTEPK		DAYKTFPPTEPKKDKKKKADETQALPQRQKKQQTVTLLPAADLDD
S	FNATRFASV	RFASVYAWNRKRISN	FPNITNLCPFGEVFNATRFASVYAWNRKRISNCVA
S	VRFPNITNL		VRFPNITNLCPFGEVFNATRFASVYAWN
S	HWFVTQRNF		RNFYEPQIITTD
S	YQGVNCTEV		CTEVPVAIHADQLTPT
S	HVSGTNGTK		FSNVTWFHAIHVSGTNGTKRFDN
S	WTAGAAAYY		TPGDSSSGWTA
S	TLADAGFIK		LADAGFIKQYGDCLG
S		CTFEYVSQPFLMDLE	DLEGKQGNFKN
S		TWRVYSTGSNVFQTR	
S		RVYSTGSNVFQTRAG	TGSNVFQ
S		GINITRFQTLLALHR	NITRFQ

coverage in South America was 48.29 % (0.00 %) for $S^{326-345}$ ($S^{442-456}$, $S^{1213-1227}$, and M^{61-75}). Similarly, the population coverage method was used to detect 129 MHC-II -binding allele epitopes. Despite reasonable antigenicity and immunogenicity values, 45 achieved the highest population coverage of more than 90 % (Table 10). Of note, $S^{795-809}$ and $S^{165-209}$ were predicted to be 100 % (99.53 % and 99.45 %, respectively) in North America (South America).

3.6. Identification of B-cell epitopes

B cells have the ability to recognize infectious pathogens or cancer cells and provide long-term protection by producing antibodies. These antibodies recognize antigen by binding highly selectively to an epitope. This recognition is used in subunit vaccines to provide long-term protection against the desired pathogens Jespersen et al. (2017).

The BepiPred 2.0 tool was used to predict B-cell epitopes of different lengths. A total of 34 sequential B-cell epitopes were predicted from the IEDB database in proteins of SARS-CoV-2. These B-cell epitopes were listed based on their position, sequence, length, and conservation in Table 11. It is worth noting that of the total 34 epitopes ranging in length from 1 to 52 mer, the best epitopes of protein S, namely $S^{748-748}$, $S^{602-606}$, $S^{555-562}$, and $S^{1107-1118}$, had a conservation of 100 %, 99.72

Table 13
CTL and HTL epitopes used to produce a multiepitope vaccine; their original strain and their conservation in other strain variants.

CTL Epitopes			
Protein	Sequence	Initial sequence prediction	Subtypes
N	SSPDDQIGY	Original Strain	
	SPDDQIGYY	Original Strain	
	KTFPPTEPK	Original Strain	
S	FNATRFASV	Original Strain	Alpha/Beta/Delta/ Gamma/Omicron
	VRFPNITNL	Original Strain	Alpha/Beta/Delta/ Gamma
	HWFVTQRNF	Delta	Alpha/Beta/Gamma/ Omicron
	YQGVNCTEV	Delta	Alpha/Beta/Gamma/ Omicron
	HVSGTNGTK	Original Strain	Alpha/Beta/Delta/ Gamma
	WTAGAAAYY	Original Strain	Alpha/Beta/Delta/ Gamma/Omicron
	TLADAGFIK	Original Strain	Alpha/Beta/Delta/ Gamma/Omicron
HTL Epitopes			
Protein	Sequence	Initial sequence prediction	Subtypes
M	SELVIGAVILRGHLR	Original Strain	
N	QIGYYRRATRRIRGG	Original Strain	
	IGYYRRATRRIRGGD	Original Strain	
S	CTFEYVSQPFLMDLE	Original Strain	Alpha/Beta/Delta/ Gamma/Omicron
	RFASVYAWNRKRISN	Original Strain	Alpha/Beta/Delta/ Gamma/Omicron
	TWRVYSTGSNVFQTR	Original Strain	Alpha/Beta/Delta/ Gamma/Omicron
	RVYSTGSNVFQTRAG	Original Strain	Alpha/Beta/Delta/ Gamma/Omicron
	GINITRFQTLLALHR	Original Strain	Alpha/Beta/Delta/ Gamma/Omicron

%, 99.17 %, and 99.17 %, respectively. For proteins M and N, 5 and 11 (1 and 4) linear B-cell epitopes ranging from 2 to 20 mer in length had a conservation index greater (less) than 90 %. Finally, of the 2 epitopes in protein E, E^{6-9} was predicted to have a conservation index of 93.75 % and the other was discarded with a conservation index of 0 % (Table 12).

We found that 10 CTL epitopes of the original strain overlapped with B-cell epitopes and matched those found in the sequences of the Alpha, Beta, Delta, Gamma, and Omicron variants. We conclude that these epitopes in these variants are capable of activating humoral and cellular immune responses simultaneously and may play a role in vaccine construction. For vaccine construction, 6 HTL epitopes were selected that overlapped with B-cell epitopes and matched HTL epitopes of protein S. The overlap of peptides makes it possible to reduce the cost of producing a large number of peptides and should be used to evaluate the sensitivity of the study Lehtinen et al. (1995).

3.7. Multiepitope vaccine construction

Proper epitope selection is essential for vaccine construction using the *in silico* biological method Yin et al. (2016). In the current study, the epitopes (CTL and HTL) of the whole SARS-CoV-2 proteome were screened based on several immune filters. They should be *[i]* antigenic, immunogenic, nontoxic, and nonallergenic, *[ii]* have affinity for at least 2 HLA alleles, *[iii]* produce IFN-γ (HTL epitopes), *[iv]* contain overlapping HTL and CTL epitopes with B-cell epitopes, and *[v]* contain at least 90 % conserved regions. Furthermore, because this is a virus that increases its virulence potential after mutation events, we increased the potency of our vaccine by adding regions of high intervariant divergence. In this way, we selected a total of 18 epitopes, 10 CTL (N^{78-86}, N^{79-87}, $N^{361-369}$, $S^{341-349}$, $S^{326-334}$, $S^{1100-1108}$, $S^{611-619}$, S^{68-76} and $S^{257-265}$) and 8 HTL ($M^{136-150}$, $N^{310-324}$, $N^{311-323}$, $S^{166-180}$, $S^{346-360}$, $S^{632-646}$, $S^{634-648}$, and $S^{232-246}$) that will make up our vaccine prototype (Table 13.

We linked the 10 CTL epitopes and 8 HTL epitopes that overlap with the B-cell epitopes using the AAY (GPGPG) linker for CTL (HTL) epitopes to form the final vaccine construct. The AAY linker is a type of proteasome cleavage site that was used to manipulate the protein in favor of greater stability, folding, and expression patterns Shamriz et al. (2016); Abdellrazeq et al. (2020); Folegatti et al. (2020). The GPGPG was added because it could enhance some biological activities of the protein by increasing solubility and facilitating immunological processing of the vaccine construct Kavoosi et al. (2007). After the addition of linkers and adjuvant, the final vaccine construct was 325 amino acids long.

In addition, the adjuvant β-defensin was added at the N-terminus of the multi-epitope with the EAAAK linker. It helps elicit a high level of cellular and immunogenic humoral responses to specific antigens and enhances the stability and longevity of the vaccine Bonam et al. (2017). Indeed, β-defensins are peptides with antimicrobial, antiviral, antibacterial, and antifungal activity and have the ability to recruit antigen-presenting immune cells with MHC-I and MHC-II Lehrer and Lu (2012).

The constructed vaccine had high population coverage for countries with high morbidity and infection. Most studies on SARS-CoV-2 vaccine construction had focused only on S in predicting B- and T-cell epitopes Chukwudozie et al. (2021); Khan et al. (2021); Kar et al. (2020). In the present study, we predicted T-cell epitopes from the conserved regions of all 4 structural proteins of SARS-CoV-2 and from sequences containing the variants in question (alpha, beta, gamma, delta, and omicron). This certainly justifies the robustness of our vaccine construct against the new variants.

It is not difficult to conclude that the vaccine we developed has several advantages over single-epitope and conventional vaccines due to the following particular features: (a) it comprises multiple MHC epitopes and therefore can be recognized by multiple T/B cell receptors, (b) it contains overlapping CTL and HTL epitopes and therefore can activate both innate and adaptive immunity, (c) it comprises polyepitopes of virulent target antigens, and (d) it has an immune stimulator (adjuvant) to improve a long-lasting immune response (Jiang et al., 2017; He et al., 2018) (Table 14).

3.8. Construction, refinement and validation of tertiary structure model

During homology modeling, the primary structure for the final multiepitope subunit vaccine was submitted using Robetta and Swiss Model, resulting in 1 and 5 models. The best model in each case was refined using Galaxy Refine. Of the 20 models provided at this stage, the model with the best stereochemical and structural parameters was selected, namely GDT-HA (0.9053), RMSD (0.552), MolProbity (1.319), Clash score (4.0), Poor rotamers (1.3) and Ramachandran plot (97.8).

Table 14
Post-translational modification results of prototype vaccine.

Number of N glycosylation region	N glycosylation region located in exposed surface	O glycosylation region	Acetylation region	Acetylation region score	Phosphorylation region
6	115, 136, 163, 313	34, 199, 218, 264, 276, 294	1	0.467	5, 9, 22, 34, 35, 51, 76, 86, 110, 127, 134, 138, 142, 146, 152, 171, 199, 214, 218, 235, 237, 254, 256, 264, 271, 276, 294, 319

	Model	GDT-HA	RMSD	MolProbity	Clash score	Poor rotamers	Rama favored
SwissModel	Initial	10.000	0.000	2.399	16.1	1.3	88.2
	MODEL 1	0.8895	0.603	1.505	8.8	0.0	97.8
	MODEL 2	0.8789	0.613	1.908	10.8	2.7	97.8
	MODEL 3	0.8816	0.581	1.533	9.4	0.0	97.8
	MODEL 4	0.8895	0.668	1.599	8.8	1.3	97.8
	MODEL 5	**0.9053**	**0.552**	**1.319**	**4.0**	**1.3**	**97.8**
ROBETTA	Initial	10.000	0.000	1.679	3.8	0.4	91.3
	MODEL 1	0.9885	0.286	2.031	12.9	0.4	93.8
	MODEL 2	0.9892	0.304	2.029	15.2	0.4	95.0
	MODEL 3	0.9892	0.287	2.056	12.1	1.2	94.1
	MODEL 4	0.9854	0.294	2.076	14.4	0.8	93.8
	MODEL 5	0.9831	0.301	1.957	12.7	0.4	95.0

Fig. 2. 3D structural conformation of the multi-epitope subunit vaccine after homology modeling and refinement by SwissModel and GalaxyRefine servers.

The superimposed structure of the template and the vaccine construct along with the template scores is shown in Fig. 2. Comparison of the quality indicators shows that the original tertiary structure models improved after refinement.

Ramachandran plot, Z-score, ERRAT, and Verify3D analyzes were performed to verify the structural quality of the predicted model. After (before) refinement, 97.33 % (92.77 %) of the structure was in the preferred region of the plot, whereas 2.67 % (6.60 %) of the residuals were in the allowed region and 0.00 % (0.63 %) of the structure was in the outlier regions, demonstrating the overall quality of the vaccine construct (Fig. 3).

The Z-score of − 0.88, which was in the range of scores of proteins of comparable size, indicates the reliability of the predicted model (3a) Wiederstein and Sippl (2007). The ERRAT score was 60.4651, significantly higher than the threshold of 50 Messaoudi et al. (2013) (3b). We also use the Structure Assessment server for Ramachandran plot analysis and Local Quality Estimation, which produces results consistent with GalaxyRefine results. After refinement, the model showed an overall QMEANDisco value of 0.35 + /- 0.05 Fig. 4.

3.9. Physiochemical properties and solubility prediction

Physiochemical property evaluation was required for vaccine formulation to demonstrate that the developed vaccine candidate was stable and met standards. It contained a total of 325 amino acids, and its molecular weight and instability index were 35,804.75 and 42.95, respectively. The hydropathy value of the final vaccine is predicted to be − 0.332, which means that our final vaccine is hydrophilic in nature.

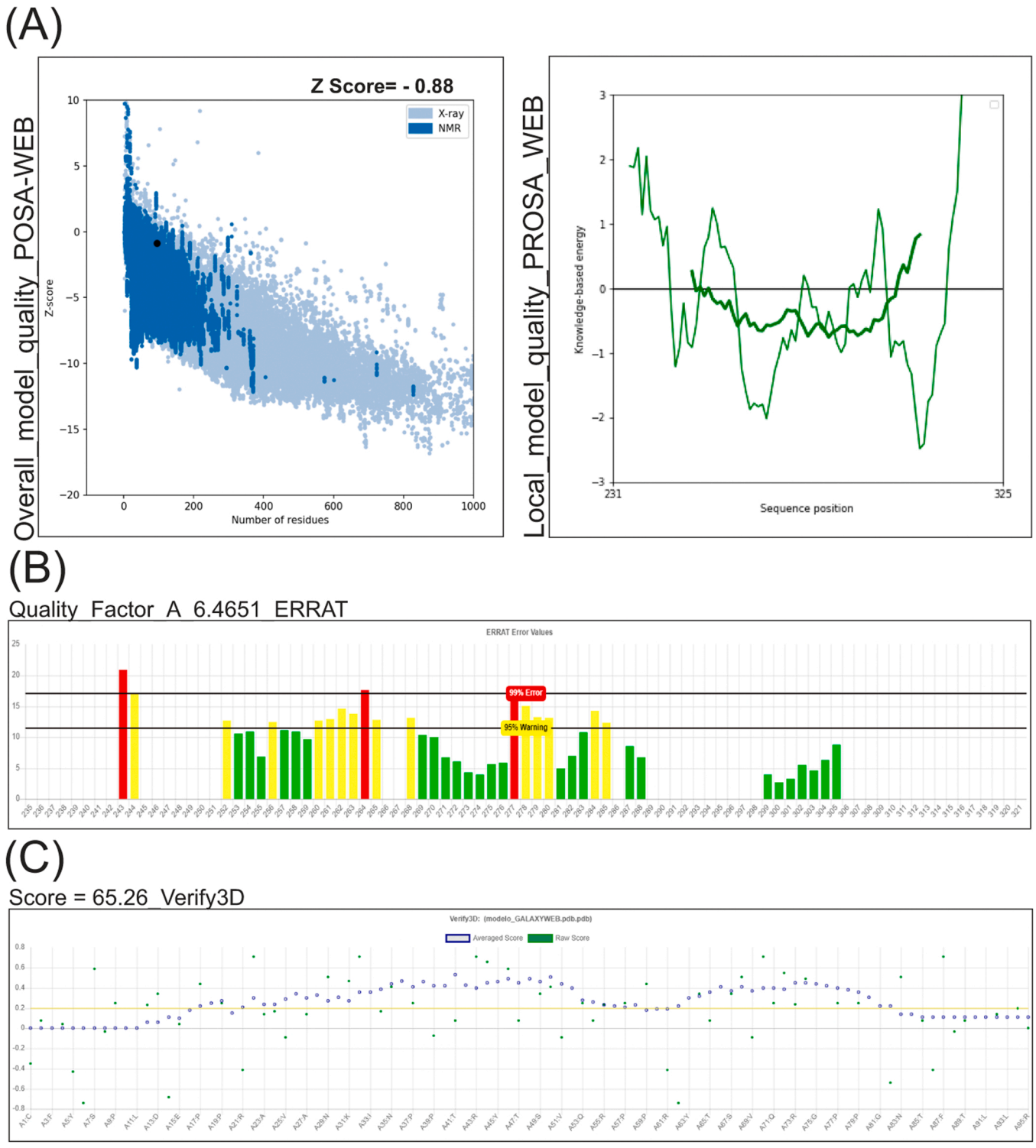


Fig. 3. Validation of the final structural subunit vaccine model (A) Vaccine 3D Structure Validation by ProSA-web illustrating Z-score; (B) Quality factor and quality score by ERRAT (C) Verify3D tools, respectively.

The half-life is predicted to be 30 h *in vitro* and > 20 h *in vivo*. The pI value of the final vaccine is calculated to be 10.04, which is an alkaline value of the strongly basic in nature.

3.10. *Prediction of post-translational modifications*

Post-translational modification analysis was performed for the prototype vaccine. The analysis revealed the presence of N-linked glycosylation positions, which is one of the major post translational modifications. It was predicted that most N-linked glycosylation positions are located on the exposed surface of the protein. This parameter is known to increase the accuracy of glycosylation Hamby and Hirst (2008). The presence of post-translational modifications in eukaryotic cells such as parasites, including *T. gondii*, is critical for selecting the correct expression system for recombinant protein production. Accordingly, the results of our post translational modifications indicated that

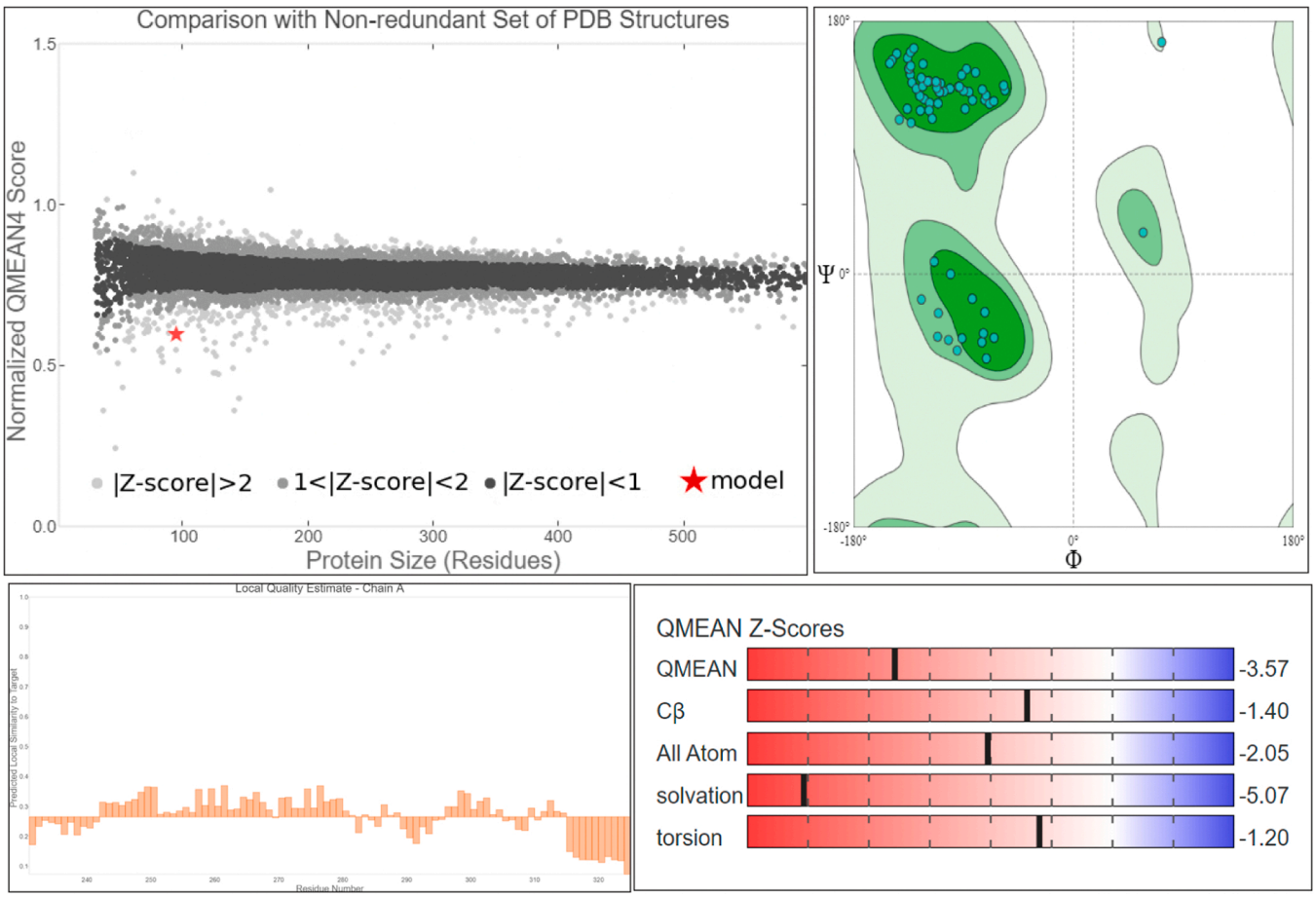


Fig. 4. Ramachandran plot indicated a high proportion of residues.

the presence of N-linked glycosylation, phosphorylation, and acetylation should be considered in the recombinant production of a vaccine prototype and that eukaryotic expression systems such as yeast, insect, and mammalian are preferable to bacterial systems (Fig. 5).

3.11. Codon optimization and E. coli expression

In silico cloning was performed so that the vaccine candidate could be expressed into the *E. coli* and the expression system was optimized by codon adaptation, which can increase protein expression up to > 1000-fold Mauro (2018). Therefore, it was necessary to optimize the codon for the vaccine construct according to the use of the *E. coli* expression system to ensure efficient translation and increased protein production Rosano and Ceccarelli (2014). The results of codon optimization of the vaccine with a GC content of 50.73 %, which is in the optimal range (30–70 %), indicate good comprehensive stability of the mRNA of the synthetic gene Grote et al. (2005); Nouri et al. (2016). The CAI value was predicted to be 1.0, in the range (0.8–1.0) signifies the maximum codon affinity and indicates high expression of our designed vaccine constructs in *E. coli* K12 strain Chen (2012). The generated cDNA sequence was 1, 263 nucleotides long after codon optimization. Finally, the recombinant plasmid was designed by computationally inserting the aligned codon sequences into the pET-28a (+) vector using SnapGene software Fig. 6. This research was conducted to develop a successful cloning approach that has a great chance of producing a high quality vaccine. Apart from all these analyzes and conclusions, further experimental validations are required to confirm the safety and efficacy of the designed vaccine.

3.12. Immune simulation of the vaccine

Immune simulations suggest that the vaccine elicits a consistent immune response. After each administration and repeated exposure, there was a significant increase in antibody response with a concomitant decrease in antigen levels. Administration of the vaccine by three injections was good at inducing different immunoglobulins. The humoral response was predominantly IgM > IgG, indicating some degree of seroconversion (Fig. 7a). The primary response was reflected by an increase in IgM levels that led to a decrease in antigen concentration. The secondary and tertiary responses with high immunoglobulin activities (IgG1, IgG2, IgG + IgM) were higher than the primary response (IgM). This indicates the emergence of an immune memory and thus increased antigen excretion during subsequent exposures.

In addition, multiple long-lived B cell isotypes were found, suggesting that long-lived B cells have the ability to switch isotype and develop memory cells (Fig. 7b). This indicates a high number of TH cells and consequently efficient Ig production supporting a humoral response (7c), as the increased TH cells promote clonal growth of B cells and antibody production Smith et al. (2000).

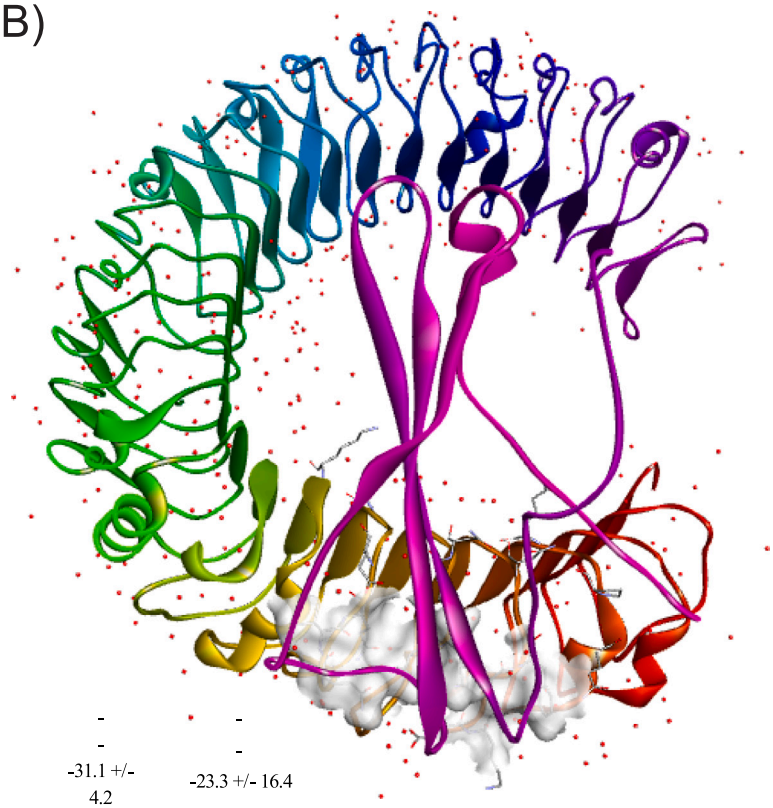
In cell populations TH (helper cells) and TC (cytotoxic cells) with corresponding memory development, the results were much closer to memory development (7d). A similar enhanced response was observed, indicating the immunogenicity of the T cell epitopes included in the vaccine construct. TC cells increased to near the maximum of over 1160 cells per mm^3 (Fig. 7e). NK cell numbers were also increased, averaging 350 cells per mm^3 (Fig. 7f). This is observed in the preactivation of the TC cell response during vaccination.

The activity of NK cells (Fig. 7g) and dendritic cells was noted along with higher macrophage activity during exposure. The number of active

(A)

RANK	Global Energy	Attractive VdW	Repulsive VdW	ACE	HB
1	**-12.22**	**-21.57**	**8.66**	**14.06**	**-6.91**
2	-10.31	-30.57	20.53	8.29	-3.81
3	21.79	-30.59	31.79	21.47	-5.27
4	25.04	-3.01	0.44	4.04	0.00
5	25.08	-12.10	4.93	9.19	-0.27
6	30.24	-12.35	21.82	10.55	-2.49
7	52.98	-24.67	83.43	6.81	-4.77
8	153.46	-8.07	205.60	-5.47	0.00
9	819.73	-16.91	1050.14	-0.84	-1.12
10	2304.68	-31.71	2893.91	14.18	-7.33

(B)



Vaccine -TLR3	Van der Waals energy - E_{vdw} (kcal/mol-1)	Electrostatic energy - E_{elec} (kcal/mol-1)	Desolvation energy (kcal/mol-1)	ΔG (kcal mol-1)	Kd (M) at 25.0 °C	RMSD (Å) to PatchDock
PatchDock	-	-	-	-8.4	2.6E-08	-
FireDock	-	-	-	-8.9	9.6E-08	1.14
QM/MM	-31.1 +/-3.2	-23.3 +/-3.4	-1.1 +/-3.3	-16.5	6.7E-07	3.90

- - -31.1 +/- 4.2 - - -23.3 +/- 16.4

Fig. 5. Structure and refinement of the docking complex. (A) Refinement of the docking complex using the FireDock tool. (B) Docked vaccine TLR3 complex and binding affinities before and after QM/MM simulation.

macrophages increased with subsequent doses and then decreased to less than 25 cells per mm3 days after the third dose (7h). This suggests another good performance indicator of the vaccine construct and demonstrates its ability to stimulate the correct immunologic compartment for an effective response, considering that IFN-γ is released from NK and the presence of IFN-γ stimulates macrophage activation Tau and Rothman (1999).

High levels of IFN-γ and IL − 2 were also detected, consistent with the prediction of IFN-γ epitopes in the vaccine. Because it contains antiviral components, IFN-γ plays an important role in both innate and adaptive immunity. A significant increase in the levels of IFN-γ, IL − 10, IL − 23, and IL − 12 was also observed with subsequent exposure. The increased IFN-γ production justifies the selection of IFN-γ producing MHC class II epitopes. This implies that the vaccine elicits a robust immune response with a short exposure and that immunity also increases with subsequent, repeated exposure (Fig. 7i).

Cytokine and interleukin production is also indicative of a successful immune response, as IFN-γ and IL − 2 production remained constant

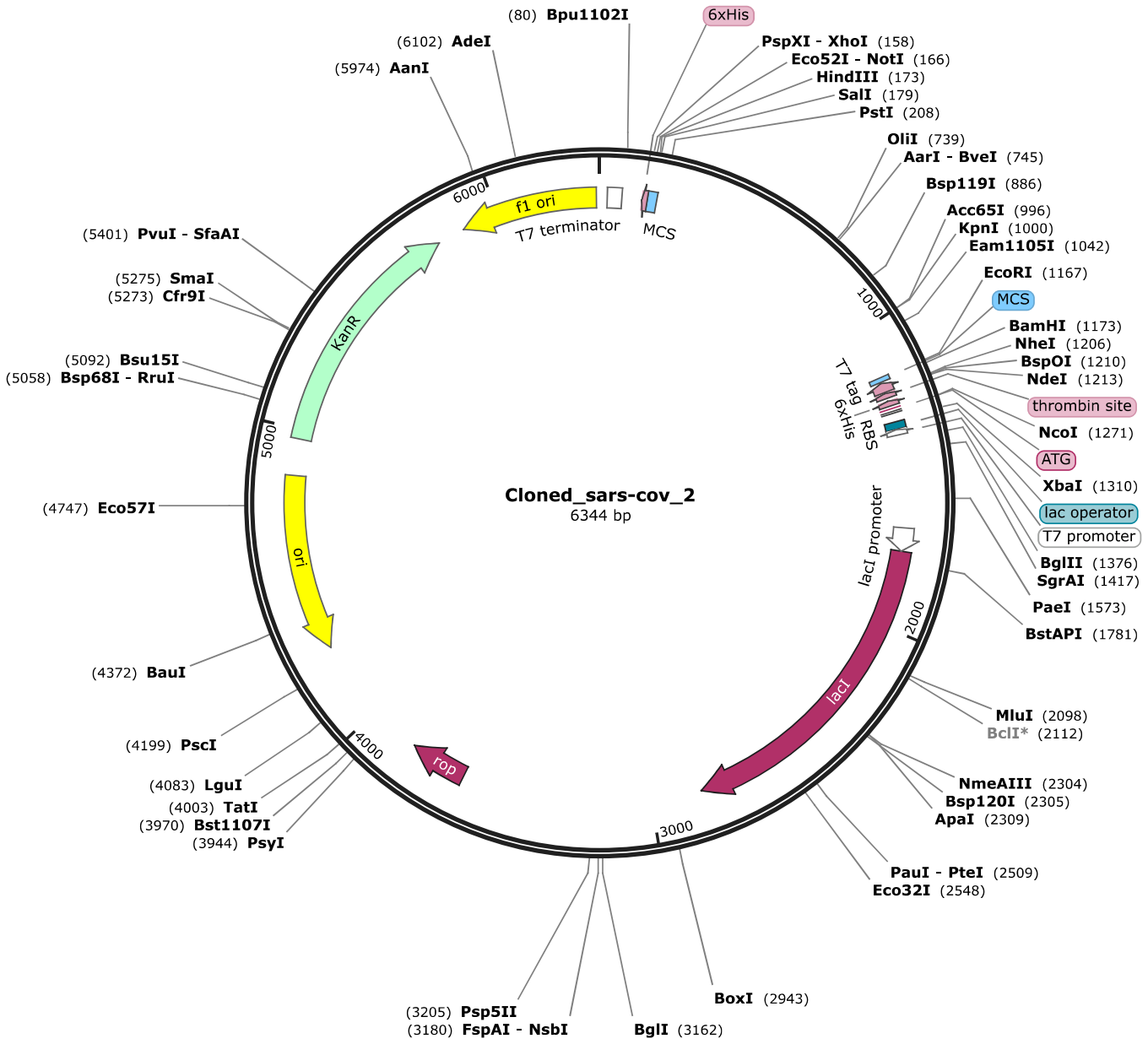


Fig. 6. In silico restriction cloning of the final vaccine construct into pET28a (+) expression vector where red part representing the vaccine insert and black circle showing the vector.

after the first injection, consistent with the prediction of IFN-γ epitopes in the vaccine. Significant increases in IFN-γ levels IL − 10, IL − 23, and IL − 12, which are important for co-stimulatory signaling of T-cell activation, were also observed during subsequent exposure. This suggests that the vaccine elicits a robust immune response during a brief exposure and that immunity also increases during subsequent, repeated exposures.

In addition, after the first dose, there was an initial increase in IFN-g responses (associated with both CD8 + T cells and CD4 + Th 1 responses) and IL − 10 & TGF-b cytokines, which are associated with the T-reg phenotype. These immune responses are also associated with a broader base in peak antigen quantification. However, after the first booster dose, IFNg also peaked. At the second booster dose on day 56, the IgG response was faster and higher than the IgM response (suggesting complete seroconversion and a B-cell memory response), faster antigen clearance as evidenced by a narrower base of the antigen spike compared with the previous doses, and lower levels of T-reg-associated cytokines (TGFb & IL −10) compared with the response to the first booster dose. Overall, the results showed a concomitant increase in immune response with each vaccination regimen.

Further studies of immunogenicity, efficacy, and possible adverse effects should be conducted both *in vitro* and *in vivo*, including expression of this vaccine candidate in a bacterial system to verify immunoreactivity by serologic analysis.

3.13. Molecular docking, QM/MM study and binding profile

Interaction between the prototype vaccine and an appropriate immune receptor molecule is required for adequate elucidation of the immune response. Antiviral immunity is mainly activated by the family of cytosolic receptors for pathogen recognition, including TLR3. Activation of TLR3 impairs replication of several viruses, including SARS-CoV Gralinski et al. (2017), MERS-CoV Mubarak et al. (2019), dengue virus (DENV), human immunodeficiency virus (HIV), influenza virus, and respiratory syncytial virus (RSV), herpes simplex virus (HSV), and Marek's disease virus (MDV) Perales-Linares and Navas-Martin (2013);

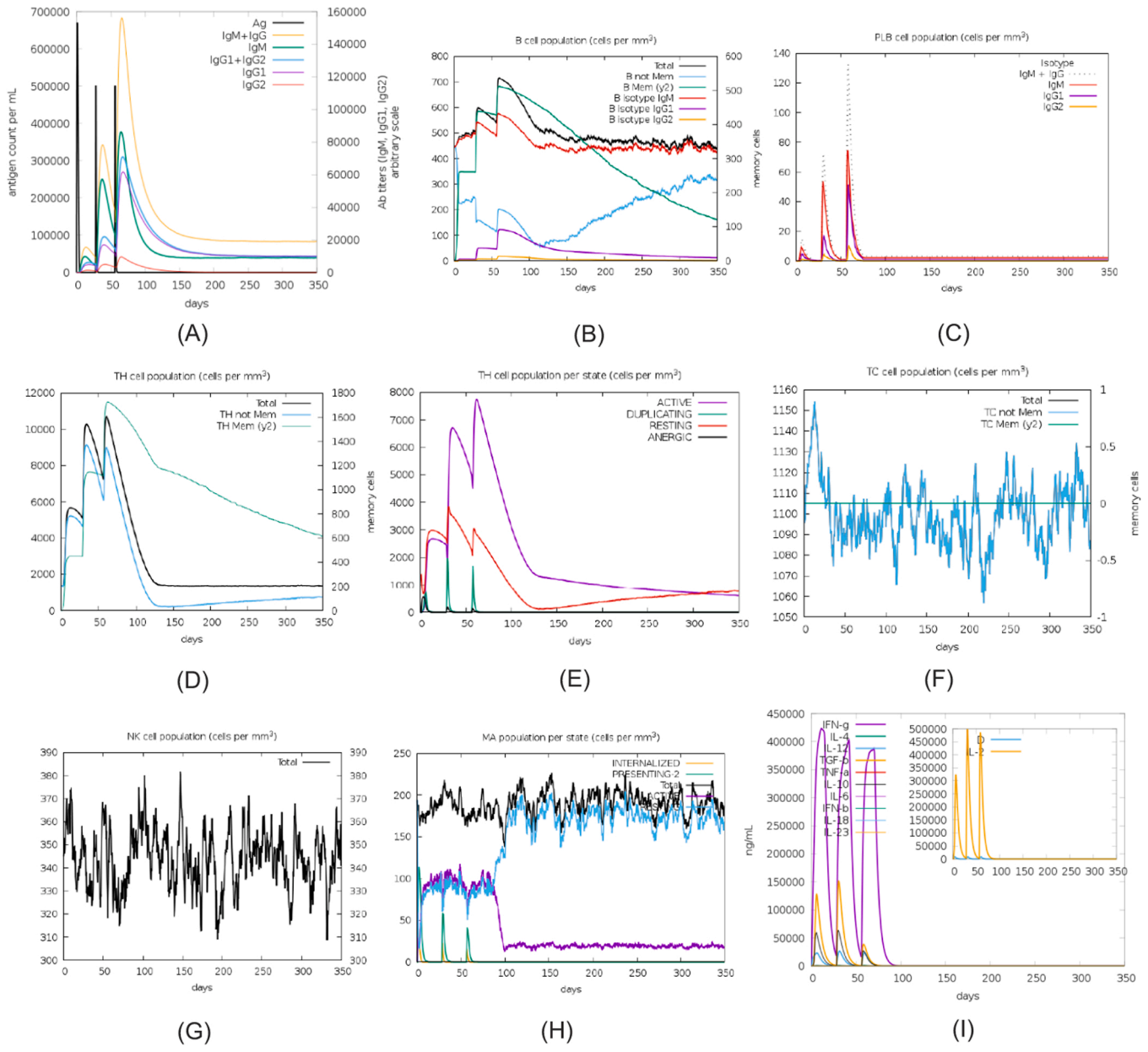


Fig. 7. C-ImmSim server prediction results of immune response after administering vaccine construct; (a) Ab titer increases with each successive injection as antigen number decreases; (b) B cell population (indicates increase in various B cell types and their potential to switch classes); (c) PLB (plasma B cell) population (indicates presence of memory B cells and active B cell proliferation); (d) TH cell population (indicates significant increase in TH memory cells); (E) TH cell population per state (indicates increase in TH cells in active state); (F) TC cell population (indicates fluctuation of TC cell population with time); (G) NK cell population (indicates the fluctuation of NK cell population with time); (H) macrophage population (indicates the fluctuation of macrophage population with time); I) concentration of cytokines and interleukins (indicates the increased IFN-γ and IL-2 production).

Prathyusha et al. (2020).

Here, the X-ray data of the extracellular domain of Human Toll-like Receptor 3 with a resolution of 2.10 Å (PDB ID: 1ZIW) were used as a template for the docking experiments with our vaccine prototype. Vaccine-TLR3 docking studies were performed and refined by PatchDock and FireDock simulations, respectively. The Firedock score of − 12.22 for global binding energy indicates good binding affinity, and a negative score indicates better docking, in addition to adequate values for attractive (−21.57) and repulsive (8.66) van der Walls forces, desolvation (2.23), and HB (−6.91) energies (Fig. 5a).

It is known that due to the low accuracy of the scoring functions, more robust approaches should be used to find the best pose among the docking results Cho et al. (2009). Another important issue related to ligand-protein docking is to consider the flexibility of the protein structure to make the results more realistic Brooijmans and Kuntz (2003); Burger et al. (2011). Therefore, quantum mechanical/molecular mechanical (QM /MM) calculations were subsequently performed to optimize the molecular geometries within the framework of the density functional theory formalism (DFT) (Fig. 5b) to allow for some degree of flexibility in the binding pocket of the receptor, which is essential for its function and to correct for vaccine matching but is not present in the crude docking procedure Sousa et al. (2006); de Medeiros et al. (2016); Harmalkar and Gray (2021). The efficacy of this procedure has already been validated in the development of improved schizophrenia drugs Zanatta et al. (2014) and multiepitope peptide vaccines against Mayaro virus Silva et al. (2021); da Silva et al. (2022).

The improved vaccine-TLR3 structure had a binding energy (ΔG) and dissociation constant (Kd) of − 16.5 kcal/mol and 6.7E-07, respectively,

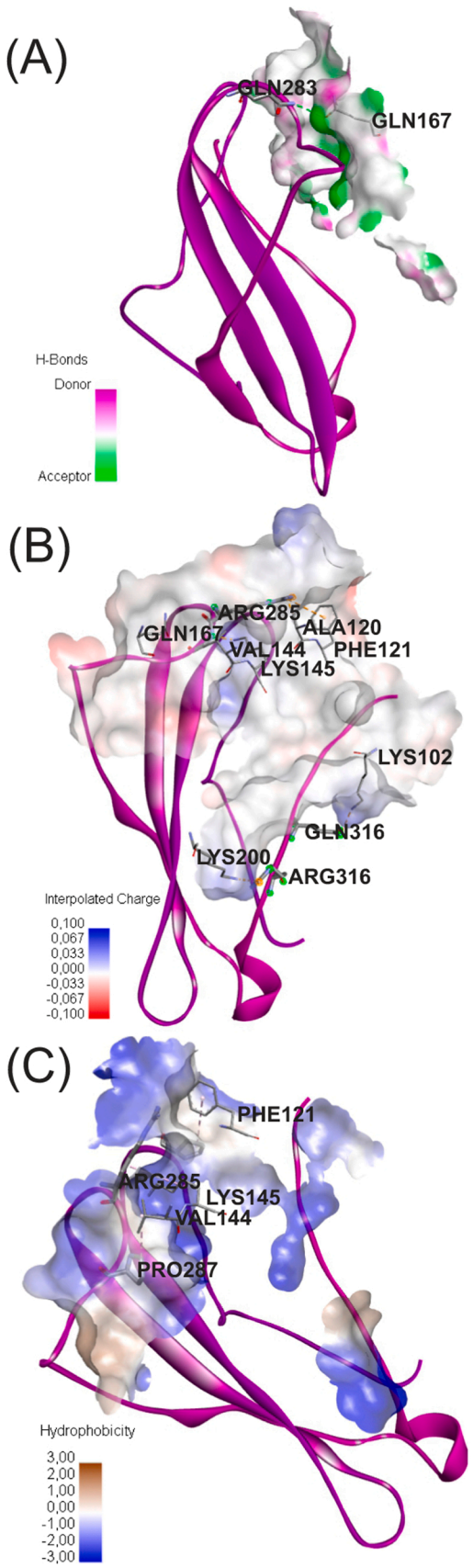


Fig. 8. Docking complex exhibiting intermolecular interactions between the vaccine component and TLR − 3, where in the receptor surface is characterized by (A) H-bond donor-acceptor, (B) interpolated charge, and (C) hydrophobicity.

at 25 ˚C, whereas the first structure ΔG was equal to − 8.4 and Kd of 2.6E-8. Moreover, there are 322 intermolecular contacts (ICs) at the interface within the threshold distance of 5.5 $\AA$, and the percentage of charged and apolar noninteracting surface (NIS %) is 33.12 % and 41.28 %, respectively. The van der Waals energy (E_{vdw}), electrostatic energy (E_{elec}), and desolvation energy were − 31.1 kcal/mol, − 23.3 kcal/mol, and − 1.1 kcal/mol, respectively. RMSD analysis revealed a deviation of 3.9 for the simulated QM:MM vaccine-TLR3 complex from the original structure.

Finally, the QM/MM simulation shows important binding contacts, namely GLN283-GLN167, ARG285-PHE121, ALA120-ARG285, LYS145-GLY286, LYS200-ARG316, LYS102-GLN318, VAL144-PRO287, ARG285-LYS145, and PHE121-PRO249 in the vaccine-TLR3 system (Fig. 8). GLN283-GLN167 is characterized by non-classical hydrogen bonding, indicating that the donor is a polarized carbon atom. The cationic part of the guanidinium of ARG285 and the anionic carboxylate group of PHE121 have electrostatic attraction. ALA120-ARG285, LYS145-GLY286, LYS200-ARG316, and LYS102-GLN318 also exert electrostatic attraction. VAL144-PRO287 and ARG285-LYS145 belong to the hydrophobic alkyl type and have a surface area equal to or greater than the area of a methyl group multiplied by the surface area scaling factor (default value 0.65), which defines them as non-polarized, non-P. In contrast, PHE121-PRO249 are characterized as a hydrophobic Pi/alkyl mixture, defined as weak interactions between a hydrogen and a Pi ring system.

Considering that TLR-3 has been shown to have a recognition function in both SARS-CoV and MERS-CoV Gralinski et al. (2017); Mubarak et al. (2019) and SARS-CoV-2 has a similar genome organization to these, it is hypothesized that this immune receptor is involved in the response against the current pandemic virus. The pattern of molecular interactions of the vaccine with TLR-3 and our *in silico* simulations indicate that this receptor is capable of acting as a sensor for recognizing molecular patterns of the pathogen and eliciting both innate and adaptive immune responses.

When developing a multiepitope vaccine, it is important to consider the viral subtypes so that this vaccine has a higher specificity. Here, our multiepitope vaccine was tested against the BA.1 and BA.2 subtypes of Omicron (B.1.1.529) using in silico techniques, but we believe that it is also effective against other Omicron sublineages such as B.A.3 and B. A.4/5. We compared the mutational profiles of the epitopes and found that 44.5 % of the epitopes of the vaccine prototype are unique to the Omicron subvariants, *i.e.*, they represent regions of the viral proteins that mutated during the emergence of the subvariants BA.1, BA.2, BA.3 and BA.4/5. More importantly, 55.5 % have completely identical peptide sequences conserved in all strains reported up toJuly 2022, namely: 5 CTL ($S^{341-349}$, $S^{1100-1108}$, $S^{611-619}$, $S^{257-265}$, and $S^{826-834}$) and 5 HTL ($S^{166-180}$, $S^{346-360}$, $S^{632-646}$, $S^{634-648}$, and $S^{232-246}$).

Immunoinformatics is one of the areas accelerating the progress of immunological research towards the development of effective vaccines. Consistent with the scientific literature, peptide epitope-based vaccines have potential outcomes against several highly infectious diseases such as SARS-COV − 2 Bhattacharya et al. (2020), MERS-COV UlQamar et al. (2019), Dengue Brinton (2002), Chikungunya Narula et al. (2018), Japanese Encephalitis Chakraborty et al. (2020), HIV Abdulla et al. (2019); Jardine et al. (2013), and Tuberculosis Ong et al. (2020). Thus, the epitopes described in this study and the prototype vaccine could be tested as diagnostic reagents and for their potential immunizing capacity against SARS-CoV-2, respectively.

4. Conclusion

Here, we developed a multiple epitope-based subunit vaccine against the circulating SARS-CoV-2 variants, namely alpha, beta, gamma, delta, and omicron. Initially, we predicted B-cell and T-cell epitopes of four SARS-CoV-2 viral proteins (S, M, N, and E) obtained from 475 genomes sequenced from the regions with the highest number of registered cases

using antigenicity, immunogenicity, allergenicity, toxicity, IFN-γ inducing, population coverage, and conservation filters. A total of 18 effective epitopes were combined by using appropriate linkers and adjuvant to enhance their immunogenicity, resulted in a multi-epitope vaccine with length of 325aa. The specificity of binding of this vaccine candidate to the TLR-3 immune cell receptor was evaluated by molecular docking, followed by the application of a QM/MM energy minimization strategy to improve the quality of the result. Finally, we found that GLN283-GLN167, ARG285-PHE121, ALA120-ARG285, LYS145-GLY286, LYS200-ARG316,LYS102-GLN318, VAL144-PRO287, ARG285-LYS145 and PHE121-PRO249 formed the binding pocket of the complex.

This is the first study that has considered such a large amount of protein data, specifically 475 proteomes of the major circulating SARS-CoV-2 variants, including 38 (16) reported non-synonymous mutations found in the protein spike (nucleocapsid phosphoprotein). It is also the only multi-epitope vaccine prototype consisting of B-cell and T-cell epitopes of all structural polyproteins of this virus, developed using hybrid quantum mechanical/molecular mechanical approaches. Further *in vitro* and *in vivo* studies are required to confirm the efficacy of our vaccine prototype against the major SARS-CoV-2 variants.

CRediT authorship contribution statement

Daniel M. de O. Campos: Data curation, Investigation, Formal analysis, Visualization, Writing – original draft. **Maria K. Silva:** Formal analysis, Validation, Visualization. **Emmanuel D. Barbosa:** Formal analysis, Writing – review & editing. **Chiuan Y. Leow:** Formal analysis, Writing – review & editing. **Umberto L. Fulco:** Supervision, Formal analysis, Writing – review & editing. **Jonas I. N. Oliveira:** Conceptualization, Methodology, Funding acquisition, Project administration, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work has received financial support from the Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq, and the Coordenaçõ de Aperfeiçoamento de Pessoal de Nível Superior - CAPES. We would like to thank the Núcleo de Processamento de Alto Desempenho of the Universidade Federal do Rio Grande do Norte - NPAD/UFRN to allow us to access their computer facilities.

References

Abbas, A.K., Lichtman, A.H., Pillai, S., 2014. Cellular and molecular immunology E-book. Elsevier Health Sciences.

Abdellrazeq, G.S., Fry, L.M., Elnaggar, M.M., Bannantine, J.P., Schneider, D.A., Chamberlin, W.M., Mahmoud, A.H., Park, K.-T., Hulubei, V., Davis, W.C., 2020. Simultaneous cognate epitope recognition by bovine cd4 and cd8 t cells is essential for primary expansion of antigen-specific cytotoxic t-cells following ex vivo stimulation with a candidate mycobacterium avium subsp. paratuberculosis peptide vaccine. Vaccine 38 (8), 2016–2025.

Abdelmageed, M.I., Abdelmoneim, A.H., Mustafa, M.I., Elfadol, N.M., Murshed, N.S., Shantier, S.W., Makhawi, A.M., 2020. Design of a multiepitope-based peptide vaccine against the e protein of human covid-19: an immunoinformatics approach. BioMed. Res. Int.

Abdulla, F., Adhikari, U.K., Uddin, M.K., 2019. Exploring t & b-cell epitopes and designing multi-epitope subunit vaccine targeting integration step of hiv-1 lifecycle using immunoinformatics approach. Microb. Pathog. 137, 103791.

AlSaba, A., Adiba, M., Saha, P., Hosen, M.I., Chakraborty, S., Nabi, A.N., 2021. An in-depth in silico and immunoinformatics approach for designing a potential multi-epitope construct for the effective development of vaccine to combat against sars-cov-2 encompassing variants of concern and interest. Comput. Biol. Med. 136, 104703.

Andreatta, M., Nielsen, M., 2018. Bioinformatics tools for the prediction of t-cell epitopes. In: Epitope Mapping Protocols. Springer, pp. 269–281.

Andreatta, M., Karosiene, E., Rasmussen, M., Stryhn, A., Buus, S., Nielsen, M., 2015. Accurate pan-specific prediction of peptide-mhc class ii binding affinity with improved binding core identification. Immunogenetics 67 (11–12), 641–650.

Andrusier, N., Nussinov, R., Wolfson, H.J., 2007. Firedock: fast interaction refinement in molecular docking. Proteins: Struct. Funct. Bioinformatics 69 (1), 139–159.

Barda, N., Dagan, N., Cohen, C., Hernán, M.A., Lipsitch, M., Kohane, I.S., Reis, B.Y., Balicer, R.D., 2021. Effectiveness of a third dose of the bnt162b2 mrna covid-19 vaccine for preventing severe outcomes in israel: an observational study. Lancet 398 (10316), 2093–2100.

Bezerra, K., Vianna, J., Neto, J.L., Oliveira, J.I.N., Albuquerque, E., Fulco, U., 2020. Interaction energies between two antiandrogenic and one androgenic agonist receptor in the presence of a t877a mutation in prostate cancer: a quantum chemistry analysis. N. J. Chem. 44 (15), 5903–5912.

Bhattacharya, M., Sharma, A.R., Patra, P., Ghosh, P., Sharma, G., Patra, B.C., Lee, S.-S., Chakraborty, C., 2020. Development of epitope-based peptide vaccine against novel coronavirus 2019 (sars-cov-2): Immunoinformatics approach. J. Med. Virol. 92 (6), 618–631.

Blom, N., Sicheritz-Pontén, T., Gupta, R., Gammeltoft, S., Brunak, S., 2004. Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. Proteomics 4 (6), 1633–1649.

Bonam, S.R., Partidos, C.D., Halmuthur, S.K.M., Muller, S., 2017. An overview of novel adjuvants designed for improving vaccine efficacy. Trends Pharmacol. Sci. 38 (9), 771–793.

Brinton, M.A., 2002. The molecular biology of west nile virus: a new invader of the western hemisphere. Annu. Rev. Microbiol. 56 (1), 371–402.

Brooijmans, N., Kuntz, I.D., 2003. Molecular recognition and docking algorithms. Annu. Rev. Biophys. Biomol. Struct. 32 (1), 335–373.

Bui, H.-H., Sidney, J., Dinh, K., Southwood, S., Newman, M.J., Sette, A., 2006. Predicting population coverage of t-cell epitope-based diagnostics and vaccines. BMC Bioinformatics 7 (1), 153.

Bui, H.-H., Sidney, J., Li, W., Fusseder, N., Sette, A., 2007. Development of an epitope conservancy analysis tool to facilitate the design of epitope-based diagnostics and vaccines. BMC Bioinformatics 8 (1), 361.

Burger, S.K., Thompson, D.C., Ayers, P.W., 2011. Quantum mechanics/molecular mechanics strategies for docking pose refinement: distinguishing between binders and decoys in cytochrome c peroxidase. J. Chem. Inf. Model. 51 (1), 93–101.

Calis, J.J., Maybeno, M., Greenbaum, J.A., Weiskopf, D., De Silva, A.D., Sette, A., Keşmir, C., Peters, B., 2013. Properties of mhc class i presented peptides that enhance immunogenicity. PLoS Comput. Biol. 9 (10), e1003266.

Campos, D., Oliveira, C., Andrade, J., Oliveira, J., 2020. Fighting covid-19. Braz. J. Biol. 80, 698–701.

Campos, D.M., Bezerra, K.S., Esmaile, S.C., Fulco, U.L., Albuquerque, E.L., Oliveira, J.I., 2020. Intermolecular interactions of cn-716 and acyl-kr-aldehyde dipeptide inhibitors against zika virus. Phys. Chem. Chem. Phys. 22 (27), 15683–15695.

Castiglione, F., Mantile, F., De Berardinis, P., Prisco, A., 2012. How the interval between prime and boost injection affects the immune response in a computational model of the immune system. Comput. Math. Methods Med.

Chabot, I., Goetghebeur, M.M., Grégoire, J.-P., 2004. The societal value of universal childhood vaccination. Vaccine 22 (15–16), 1992–2005.

Chakraborty, S., Barman, A., Deb, B., 2020. Japanese encephalitis virus: a multi-epitope loaded peptide vaccine formulation using reverse vaccinology approach. Infect. Genet. Evol. 78, 104106.

Chen, R., 2012. Bacterial expression systems for recombinant protein production: E. coli and beyond. Biotechnol. Adv. 30 (5), 1102–1107.

Chen, V.B., Arendall, W.B., Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., Richardson, D.C., 2010. Molprobity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. Sect. D: Biol. Crystallogr. 66 (1), 12–21.

Chesler, D.A., Reiss, C.S., 2002. The role of ifn-γ in immune responses to viral infections of the central nervous system. Cytokine Growth Factor Rev. 13 (6), 441–454.

Cho, A.E., Chung, J.Y., Kim, M., Park, K., 2009. Quantum mechanical scoring for protein docking. J. Chem. Phys. 131 (13), 134108.

Chukwudozie, O.S., Gray, C.M., Fagbayi, T.A., Chukwuanukwu, R.C., Oyebanji, V.O., Bankole, T.T., Adewole, R.A., Daniel, E.M., 2021. Immuno-informatics design of a multimeric epitope peptide based vaccine targeting sars-cov-2 spike glycoprotein. PLoS One 16 (3), e0248061.

Chung, L.W., Sameera, W., Ramozzi, R., Page, A.J., Hatanaka, M., Petrova, G.P., Harris, T.V., Li, X., Ke, Z., Liu, F., et al., 2015. The oniom method and its applications. Chem. Rev. 115 (12), 5678–5796.

Cui, Z., 2005. Dna vaccine. Adv. Genet. 54, 257–289.

daSilva, M.K., Azevedo, A.A.C., Campos, D.M. d.O., de Souto, J.T., Fulco, U.L., Oliveira, J.I.N., 2022. Computational vaccinology guided design of multi-epitope subunit vaccine against a neglected arbovirus of the americas. J. Biomol. Struct. Dyn. 1–18.

Dimitrov, I., Flower, D.R., Doytchinova, I., 2013. Allertop-a server for in silico prediction of allergens. BMC Bioinformatics.

Dimitrov, I., Bangov, I., Flower, D.R., Doytchinova, I., 2014. Allertop v. 2–a server for in silico prediction of allergens. J. Mol. Model. 20 (6), 2278.

Doria-Rose, N.A., Shen, X., Schmidt, S.D., O'Dell, S., McDanal, C., Feng, W., Tong, J., Eaton, A., Maglinao, M., Tang, H., et al., 2021. Booster of mRNA-1273 vaccine reduces sars-cov-2 omicron escape from neutralizing antibodies. medRxiv.

Doytchinova, I.A., Flower, D.R., 2007. Identifying candidate subunit vaccines using an alignment-independent method based on principal amino acid properties. Vaccine 25 (5), 856–866.

Doytchinova, I.A., Flower, D.R., 2007. Vaxijen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. BMC Bioinformatics 8 (1), 1–7.
Ehreth, J., 2003. The global value of vaccination. Vaccine 21 (7–8), 596–600.
Elliott, S.L., Suhrbier, A., Miles, J.J., Lawrence, G., Pye, S.J., Le, T.T., Rosenstengel, A., Nguyen, T., Allworth, A., Burrows, S.R., et al., 2008. Phase i trial of a cd8. t-cell peptide epitope-based vaccine for infectious mononucleosis. J. Virol. 82 (3), 1448–1457.
England P.H. 2021. Sars-cov-2 variants of concern and variants under investigation in england, technical briefing 12 (2021).
Fiolet, T., Kherabi, Y., MacDonald, C.-J., Ghosn, J., Peiffer-Smadja, N., 2021. Comparing covid-19 vaccines for their characteristics, efficacy and effectiveness against sars-cov-2 and variants of concern: a narrative review. Clin. Microbiol. Infect.
Fleri, W., Paul, S., Dhanda, S.K., Mahajan, S., Xu, X., Peters, B., Sette, A., 2017. The immune epitope database and analysis resource in epitope discovery and synthetic vaccine design. Front. Immunol. 8, 278.
Folegatti, P.M., Ewer, K.J., Aley, P.K., Angus, B., Becker, S., Belij-Rammerstorfer, S., Bellamy, D., Bibi, S., Bittaye, M., Clutterbuck, E.A., et al., 2020. Safety and immunogenicity of the chadox1 ncov-19 vaccine against sars-cov-2: a preliminary report of a phase 1/2, single-blind, randomised controlled trial. Lancet 396 (10249), 467–478.
Garcia-Beltran, W., Denis St, K., Hoelzemer, A., Lam, E., Nitido, A., Sheehan, M., Berrios, C., Ofoman, O., Chang, C., Hauser, B., et al., 2021. mRNA-based covid-19 vaccine boosters induce neutralizing immunity against sars-cov-2 omicron variant. medrxiv.
Garcia-Boronat, M., Diez-Rivero, C.M., Reinherz, E.L., Reche, P.A., 2008. Pvs: a web server for protein sequence variability analysis tuned to facilitate conserved epitope discovery. Nucleic Acids Res. 36 (suppl_2), W35–W41.
Gasteiger, E., Hoogland, C., Gattiker, A., Wilkins, M.R., Appel, R.D., Bairoch, A., et al., 2005. Protein identification and analysis tools on the expasy server. In: The Proteomics Protocols Handbook. Springer, pp. 571–607.
Gralinski, L.E., Menachery, V.D., Morgan, A.P., Totura, A.L., Beall, A., Kocher, J., Plante, J., Harrison-Shostak, D.C., Schäfer, A., Pardo-Manuel de Villena, F., et al., 2017. Allelic variation in the toll-like receptor adaptor protein ticam2 contributes to sars-coronavirus pathogenesis in mice. G3: Genes Genomes Genet. 7 (6), 1653–1663.
Grote, A., Hiller, K., Scheer, M., Münch, R., Nörtemann, B., Hempel, D.C., Jahn, D., 2005. Jcat: a novel tool to adapt codon usage of a target gene to its potential expression host. Nucleic Acids Res. 33 (suppl_2), W526–W531.
Gruhler, A., Früh, K., 2000. Control of mhc class i traffic from the endoplasmic reticulum by cellular chaperones and viral anti-chaperones. Traffic 1 (4), 306–311.
Gupta, S., Kapoor, P., Chaudhary, K., Gautam, A., Kumar, R., Raghava, G.P., Consortium, O.S.D.D., et al., 2013. In silico approach for predicting toxicity of peptides and proteins. PloS One 8 (9), e73957.
Hamby, S.E., Hirst, J.D., 2008. Prediction of glycosylation sites using random forests. BMC Bioinformatics 9 (1), 1–13.
Harmalkar, A., Gray, J.J., 2021. Advances to tackle backbone flexibility in protein docking. Curr. Opin. Struct. Biol. 67, 178–186.
Harvey, W.T., Carabelli, A.M., Jackson, B., Gupta, R.K., Thomson, E.C., Harrison, E.M., Ludden, C., Reeve, R., Rambaut, A., Peacock, S.J., et al., 2021. Sars-cov-2 variants, spike mutations and immune escape. Nat. Rev. Microbiol. 19 (7), 409–424.
He, R., Yang, X., Liu, C., Chen, X., Wang, L., Xiao, M., Ye, J., Wu, Y., Ye, L., 2018. Efficient control of chronic lcmv infection by a cd4 t cell epitope-based heterologous prime-boost vaccination in a murine model. Cell. Mol. Immunol. 15 (9), 815–826.
Jardine, J., Julien, J.-P., Menis, S., Ota, T., Kalyuzhniy, O., McGuire, A., Sok, D., Huang, P.-S., MacPherson, S., Jones, M., et al., 2013. Rational hiv immunogen design to target specific germline b cell receptors. Science 340 (6133), 711–716.
Jensen, K.K., Andreatta, M., Marcatili, P., Buus, S., Greenbaum, J.A., Yan, Z., Sette, A., Peters, B., Nielsen, M., 2018. Improved methods for predicting peptide binding affinity to mhc class ii molecules. Immunology 154 (3), 394–406.
Jespersen, M.C., Peters, B., Nielsen, M., Marcatili, P., 2017. Bepipred-2.0: improving sequence-based b-cell epitope prediction using conformational epitopes. Nucleic Acids Res. 45 (W1), W24–W29.
Jiang, P., Cai, Y., Chen, J., Ye, X., Mao, S., Zhu, S., Xue, X., Chen, S., Zhang, L., 2017. Evaluation of tandem chlamydia trachomatis momp multi-epitopes vaccine in balb/c mice model. Vaccine 35 (23), 3096–3103.
Kadam, A., Sasidharan, S., Saudagar, P., 2020. Computational design of a potential multi-epitope subunit vaccine using immunoinformatics to fight ebola virus. Infect., Genet. Evol. 85, 104464.
Kar, T., Narsaria, U., Basak, S., Deb, D., Castiglione, F., Mueller, D.M., Srivastava, A.P., 2020. A candidate multi-epitope vaccine against sars-cov-2. Sci. Rep. 10 (1), 1–24.
Kavoosi, M., Creagh, A.L., Kilburn, D.G., Haynes, C.A., 2007. Strategy for selecting and characterizing linker peptides for cbm9-tagged fusion proteins expressed in escherichia coli. Biotechnol. Bioeng. 98 (3), 599–610.
Khan, A., Khan, S., Saleem, S., Nizam-Uddin, N., Mohammad, A., Khan, T., Ahmad, S., Arshad, M., Ali, S.S., Suleman, M., et al., 2021. Immunogenomics guided design of immunomodulatory multi-epitope subunit vaccine against the sars-cov-2 new variants, and its validation through in silico cloning and immune simulation. Comput. Biol. Med. 133, 104420.
Lagunas-Rangel, F.A., Chávez-Valencia, V., 2020. High il-6/ifn-γ ratio could be associated with severe disease in covid-19 patients. J. Med. Virol.
Laughlin, C., Schleif, A., Heilman, C.A., 2015. Addressing viral resistance through vaccines. Future Virol. 10 (8), 1011–1022.
Lee, G.R., Won, J., Heo, L., Seok, C., 2019. Galaxyrefine2: simultaneous refinement of inaccurate local regions and overall protein structure. Nucleic Acids Res. 47 (W1), W451–W455.
Lee, I.-J., Sun, C.-P., Wu, P.-Y., Lan, Y.-H., Wang, I.-H., Liu, W.-C., Tseng, S.-C., Tsung, S.-I., Chou, Y.-C., Kumari, M., et al., 2022. Omicron-Specific mRNA Vaccine Induced Potent Neutralizing Antibody Against Omicron But Not Other Sars-cov-2 Variants. BioRxiv.
Lehrer, R.I., Lu, W., 2012. α -defensins in human innate immunity. Immunol. Rev. 245 (1), 84–112.
Lehtinen, M., Hibma, M.H., Stellato, G., Kuoppala, T., Paavonen, J., 1995. Human t helper cell epitopes overlap b cell and putative cytotoxic t cell epitopes in the e2 protein of human papillomavirus type 16. Biochem. Biophys. Res. Commun. 209 (2), 541–546.
Lei, Y., Shao, J., Zhao, F., Li, Y., Lei, C., Ma, F., Chang, H., Zhang, Y., 2019. Artificially designed hepatitis b virus core particles composed of multiple epitopes of type a and o foot-and-mouth disease virus as a bivalent vaccine candidate. J. Med. Virol. 91 (12), 2142–2152.
Lennerz, V., Gross, S., Gallerani, E., Sessa, C., Mach, N., Boehm, S., Hess, D., VonBoehmer, L., Knuth, A., Ochsenbein, A.F., et al., 2014. Immunologic response to the survivin-derived multi-epitope vaccine emd640744 in patients with advanced solid tumors. Cancer Immunol. Immunother. 63 (4), 381–394.
Li, W., Joshi, M.D., Singhania, S., Ramsey, K.H., Murthy, A.K., 2014. Peptide vaccine: progress and challenges. Vaccines 2 (3), 515–536.
Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J.L., Dror, R.O., Shaw, D. E., 2010. Improved side-chain torsion potentials for the amber ff99sb protein force field. Proteins: Struct. Funct. Bioinformatics 78 (8), 1950–1958.
Ling, Y.-M., Chen, J.-Y., Guo, L., Wang, C.-Y., Tan, W.-T., Wen, Q., Zhang, S.-D., Deng, G.-H., Lin, Y., Kwok, H.F., 2017. , β -defensin 1 expression in hcv infected liver/liver cancer: an important role in protecting hcv progression and liver cancer development. Sci. Rep. 7 (1), 1–14.
Luckheeram, R.V., Zhou, R., Verma, A.D., Xia, B., 2012. Cd4. t cells: differentiation and functions. Clin. Dev. Immunol.
Lundegaard, C., Lamberth, K., Harndahl, M., Buus, S., Lund, O., Nielsen, M., 2008. Netmhc-3.0: accurate web accessible predictions of human, mouse and monkey mhc class i affinities for peptides of length 8-11. Nucleic Acids Res. 36 (suppl_2), W509–W512.
Lundegaard, C., Lund, O., Nielsen, M., 2008. Accurate approximation method for prediction of class i mhc affinities for peptides of length 8, 10 and 11 using prediction tools trained on 9mers. Bioinformatics 24 (11), 1397–1398.
María, R., Arturo, C., Alicia, J.-A., Paulina, M., Gerardo, A.-O., 2017. The impact of bioinformatics on vaccine design and development. Vaccines 2, 3–6.
Mauro, V.P., 2018. Codon optimization in the production of recombinant biotherapeutics: potential risks and considerations. BioDrugs 32 (1), 69–81.
de Medeiros, A.S., Zoppi, A., Barbosa, E.G., Oliveira, J.I., Fernandes-Pedrosa, M.F., Longhi, M.R., daSilva-Júnior, A.A., 2016. Supramolecular aggregates of oligosaccharides with co-solvents in ternary systems for the solubilizing approach of triamcinolone. Carbohydr. Polym. 151, 1040–1051.
Messaoudi, A., Belguith, H., Hamida, J.B., 2013. Homology modeling and virtual screening approaches to identify potent inhibitors of veb-1 β -lactamase. Theor. Biol. Med. Model. 10 (1), 22.
Mittal, A., Sasidharan, S., Raj, S., Balaji, S., Saudagar, P., 2020. Exploring the zika genome to design a potential multiepitope vaccine using an immunoinformatics approach. Int. J. Pept. Res. Ther. 1–10.
Mubarak, A., Alturaiki, W., Hemida, M.G., 2019. Middle east respiratory syndrome coronavirus (mers-cov): infection, immunological response, and vaccine development. J. Immunol. Res.
Narula, A., Pandey, R.K., Khatoon, N., Mishra, A., Prajapati, V.K., 2018. Excavating chikungunya genome to design b and t cell multi-epitope subunit vaccine using comprehensive immunoinformatics approach to control chikungunya infection. Infect., Genet. Evol. 61, 4–15.
Nielsen, M., Lundegaard, C., Worning, P., Lauemøller, S.L., Lamberth, K., Buus, S., Brunak, S., Lund, O., 2003. Reliable prediction of t-cell epitopes using neural networks with novel sequence representations. Protein Sci. 12 (5), 1007–1017.
Nielsen, M., Lundegaard, C., Blicher, T., Lamberth, K., Harndahl, M., Justesen, S., Røder, G., Peters, B., Sette, A., Lund, O., et al., 2007. Netmhcpan, a method for quantitative predictions of peptide binding to any hla-a and-b locus protein of known sequence. PloS One 2 (8).
Nielsen, P.H., Kragelund, C., Seviour, R.J., Nielsen, J.L., 2009. Identity and ecophysiology of filamentous bacteria in activated sludge. FEMS Microbiol. Rev. 33 (6), 969–998.
Nouri, H.R., Karkhah, A., Varasteh, A., Sankian, M., 2016. Expression of a chimeric allergen with high rare codons content in codon bias-adjusted Escherichia coli: escherichia coli bl21 (de3)-codon plus ril as an efficient host. Curr. Microbiol. 73 (1), 91–98.
de Oliveira Campos, D.M., Fulco, U.L., de Oliveira, C.B.S., Oliveira, J.I.N., 2020. Sars-cov-2 virus infection: targets and antiviral pharmacological strategies. J. Evid. Based Med.
de Oliveira Campos, D.M., daSilva, M.K., Silva de Oliveira, C.B., Fulco, U.L., Nobre Oliveira, J.I., 2022. Effectiveness of covid-19 vaccines against omicron variant. Immunotherapy 0.
Ong, E., He, Y., Yang, Z., 2020. Epitope promiscuity and population coverage of mycobacterium tuberculosis protein antigens in current subunit vaccines under development. Infect. Genet. Evol., 104186
W.H. Organization, et al., Covid-19 weekly epidemiological update, edition 88, published 20 april 2022 (2022).
Perales-Linares, R., Navas-Martin, S., 2013. Toll-like receptor 3 in viral pathogenesis: friend or foe? Immunology 140 (2), 153–167.
Peters, B., Bulik, S., Tampe, R., Van Endert, P.M., Holzhütter, H.-G., 2003. Identifying mhc class i epitopes by predicting the tap transport efficiency of epitope precursors. J. Immunol. 171 (4), 1741–1749.

Ponomarenko, J.V., Bourne, P.E., 2007. Antibody-protein interactions: benchmark datasets and prediction tools evaluation. BMC Struct. Biol. 7 (1), 1–19.
Prathyusha, A., Bhukya, P.L., Bramhachari, P.V., 2020. Potentiality of toll-like receptors (tlrs) in viral infections. In: Dynamics of Immune Activation in Viral Diseases. Springer, pp. 149–159.
Rakib, A., Sami, S.A., Mimi, N.J., Chowdhury, M.M., Eva, T.A., Nainu, F., Paul, A., Shahriar, A., Tareq, A.M., Emon, N.U., et al., 2020. Immunoinformatics-guided design of an epitope-based vaccine against severe acute respiratory syndrome coronavirus 2 spike glycoprotein. Comput. Biol. Med. 124, 103967.
Rapin, N., Lund, O., Bernaschi, M., Castiglione, F., 2010. Computational immunology meets bioinformatics: the use of prediction tools for molecular binding in the simulation of the immune system. PloS One 5 (4), e9862.
Rappuoli, R., 2001. Reverse vaccinology, a genome-based approach to vaccine development. Vaccine 19 (17–19), 2688–2691.
Ren, J., Wen, L., Gao, X., Jin, C., Xue, Y., Yao, X., 2008. Css-palm 2.0: an updated software for palmitoylation sites prediction. Protein Eng. Des. Sel. 21 (11), 639–644.
Reynisson, B., Barra, C., Kaabinejadian, S., Hildebrand, W.H., Peters, B., Nielsen, M., 2020. Improved prediction of mhc ii antigen presentation through integration and motif deconvolution of mass spectrometry mhc eluted ligand data. J. Proteome Res. 19 (6), 2304–2315.
Rosano, G.L., Ceccarelli, E.A., 2014. Recombinant protein expression in Escherichia coli: advances and challenges. Front. Microbiol. 5, 172.
Scarselli, M., Giuliani, M.M., Adu-Bobie, J., Pizza, M., Rappuoli, R., 2005. The impact of genomics on vaccine design. Trends Biotechnol. 23 (2), 84–91.
Schneidman-Duhovny, D., Inbar, Y., Nussinov, R., Wolfson, H.J., 2005. Patchdock and symmdock: servers for rigid and symmetric docking. Nucleic Acids Res. 33 (suppl_2), W363–W367.
Seder, R.A., Darrah, P.A., Roederer, M., 2008. T-cell quality in memory and protection: implications for vaccine design. Nat. Rev. Immunol. 8 (4), 247–258.
Senn, H.M., Thiel, W., 2009. Qm/mm methods for biomolecular systems. Angew. Chem. Int. Ed. 48 (7), 1198–1229.
Shamriz, S., Ofoghi, H., Moazami, N., 2016. Effect of linker length and residues on the structure and stability of a fusion protein with malaria vaccine application. Comput. Biol. Med. 76, 24–29.
Silva, M.K., Gomes, H.S., Silva, O.L., Campanelli, S.E., Campos, D.M., Araújo, J.M., Fernandes, J.V., Fulco, U.L., Oliveira, J.I., 2021. Identification of promiscuous t cell epitopes on mayaro virus structural proteins using immunoinformatics, molecular modeling, and qm: Mm approaches. Infect., Genet. Evol. 91, 104826.
Smith, K.M., Pottage, L., Thomas, E.R., Leishman, A.J., Doig, T.N., Xu, D., Liew, F.Y., Garside, P., 2000. Th1 and th2 cd4. t cells provide help for b cell clonal expansion and antibody synthesis in a similar manner in vivo. J. Immunol. 165 (6), 3136–3144.
Soleymani, S., Tavassoli, A., Housaindokht, M.R., 2022. An overview of progress from empirical to rational design in modern vaccine development, with an emphasis on computational tools and immunoinformatics approaches. Comput. Biol. Med. 140, 105057.
Soria-Guerra, R.E., Nieto-Gomez, R., Govea-Alonso, D.O., Rosales-Mendoza, S., 2015. An overview of bioinformatics tools for epitope prediction: implications on vaccine development. J. Biomed. Inform. 53, 405–414.
Sousa, S.F., Fernandes, P.A., Ramos, M.J., 2006. Protein-ligand docking: current status and future challenges. Proteins: Struct. Funct. Bioinformatics 65 (1), 15–26.
Steentoft, C., Vakhrushev, S.Y., Joshi, H.J., Kong, Y., Vester-Christensen, M.B., Schjoldager, K.T.-B., Lavrsen, K., Dabelsteen, S., Pedersen, N.B., Marcos-Silva, L., et al., 2013. Precision mapping of the human o-galnac glycoproteome through simplecell technology. EMBO J. 32 (10), 1478–1488.
Stobart, C.C., Moore, M.L., 2014. Rna virus reverse genetics and vaccine design. Viruses 6 (7), 2531–2550.
Tau, G., Rothman, P., 1999. Biologic functions of the ifn-γ receptors. Allergy 54 (12), 1233.
ul Qamar, M.T., Alqahtani, S.M., Alamri, M.A., Chen, L.-L., 2020. Structural basis of sars-cov-2 3clpro and anti-covid-19 drug discovery from medicinal plants. J. Pharm. Anal. 10 (4), 313–319.
UlQamar, M.T., Saleem, S., Ashfaq, U.A., Bari, A., Anwar, F., Alqahtani, S., 2019. Epitope-based peptide vaccine design and target site depiction against middle east respiratory syndrome coronavirus: an immune-informatics study. J. Transl. Med. 17 (1), 362.
Vianna, J.F., Bezerra, K.S., Oliveira, J.I., Albuquerque, E.L., Fulco, U.L., 2019. Binding energies of the drugs capreomycin and streptomycin in complex with tuberculosis bacterial ribosome subunits. Phys. Chem. Chem. Phys. 21 (35), 19192–19200.
Vita, R., Overton, J.A., Greenbaum, J.A., Ponomarenko, J., Clark, J.D., Cantrell, J.R., Wheeler, D.K., Gabbard, J.L., Hix, D., Sette, A., et al., 2015. The immune epitope database (iedb) 3.0. Nucleic Acids Res. 43 (D1), D405–D412.
Waterhouse, A.M., Procter, J.B., Martin, D.M., Clamp, M., Barton, G.J., 2009. Jalview version 2–a multiple sequence alignment editor and analysis workbench. Bioinformatics 25 (9), 1189–1191.
WHO, Covid-19 vaccine tracker and landscape, Available at: ⟨https://www.who.int/publications/m/item/draft-landscape-of-covid-19-candidate-vaccines⟩.Accessed on 08 February 2022 (2022).
Wiederstein, M., Sippl, M.J., 2007. Prosa-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res. 35 (suppl_2), W407–W410.
Williams, C.J., Headd, J.J., Moriarty, N.W., Prisant, M.G., Videau, L.L., Deis, L.N., Verma, V., Keedy, D.A., Hintze, B.J., Chen, V.B., et al., 2018. Molprobity: more and better reference data for improved all-atom structure validation. Protein Sci. 27 (1), 293–315.
Yazdani, Z., Rafiei, A., Valadan, R., Ashrafi, H., Pasandi, M., Kardan, M., 2020. Designing a potent l1 protein-based hpv peptide vaccine: a bioinformatics approach. Comput. Biol. Chem. 85, 107209.
Yin, D., Li, L., Song, X., Li, H., Wang, J., Ju, W., Qu, X., Song, D., Liu, Y., Meng, X., et al., 2016. A novel multi-epitope recombined protein for diagnosis of human brucellosis. BMC Infect. Dis. 16 (1), 1–8.
Zanatta, G., Nunes, G., Bezerra, E.M., daCosta, R.F., Martins, A., Caetano, E.W., Freire, V. N., Gottfried, C., 2014. Antipsychotic haloperidol binding to the human dopamine d3 receptor: beyond docking through qm/mm refinement toward the design of improved schizophrenia medicines. ACS Chem. Neurosci. 5 (10), 1041–1054.
Zang, J., Zhang, C., Yin, Y., Xu, S., Qiao, W., Lavillette, D., Wang, H., Huang, Z., 2022. An mRNA Vaccine Candidate For The Sars-cov-2 Omicron Variant. BioRxiv.
Zhao, Z., Ma, X., Zhang, R., Hu, F., Zhang, T., Liu, Y., Han, M.H., You, F., Yang, Y., Zheng, W., 2021. A novel liposome-polymer hybrid nanoparticles delivering a multi-epitope self-replication dna vaccine and its preliminary immune evaluation in experimental animals. Nanomed.: Nanotechnol. Biol. Med. 35, 102338.
Zuin, M., Zuliani, G., Roncon, L., 2021. High covid-19 transmission rates jeopardize global mass vaccination campaigns. Pathog. Glob. Health 115 (4), 213–214.