

Original Paper

# Development of a Multiepitope Vaccine Against SARS-CoV-2: Immunoinformatics Study

Fatemeh Ghafouri<sup>1</sup>, BSc; Reza Ahangari Cohan<sup>2</sup>, PhD; Hilda Samimi<sup>3</sup>, PhD; Ali Hosseini Rad S M<sup>4</sup>, PhD; Mahmood Naderi<sup>5</sup>, PhD; Farshid Noorbakhsh<sup>6</sup>, MD, PhD; Vahid Haghpanah<sup>3,7</sup>, MD, MPH, PhD

<sup>1</sup>Department of Biotechnology, Faculty of Life Sciences and Biotechnology, Shahid Beheshti University, Tehran, Iran

<sup>2</sup>Department of Nanobiotechnology, New Technologies Research Group, Pasteur Institute of Iran, Tehran, Iran

<sup>3</sup>Endocrinology and Metabolism Research Center, Endocrinology and Metabolism Clinical Sciences Institute, Tehran University of Medical Sciences, Tehran, Iran

<sup>4</sup>Department of Microbiology and Immunology, University of Otago, Otago, New Zealand

<sup>5</sup>Digestive Diseases Research Center, Digestive Diseases Research Institute, Tehran University of Medical Sciences, Tehran, Iran

<sup>6</sup>Department of Immunology, School of Medicine, Tehran University of Medical Sciences, Tehran, Iran

<sup>7</sup>Personalized Medicine Research Center, Endocrinology and Metabolism Clinical Sciences Institute, Tehran University of Medical Sciences, Tehran, Iran

**Corresponding Author:**

Vahid Haghpanah, MD, MPH, PhD

Endocrinology and Metabolism Research Center

Endocrinology and Metabolism Clinical Sciences Institute

Tehran University of Medical Sciences

Dr Shariati Hospital

North Kargar Avenue

Tehran, 14114

Iran

Phone: 98 21 88220037

Email: [v.haghpanah@gmail.com](mailto:v.haghpanah@gmail.com)

## Abstract

**Background:** Since the first appearance of SARS-CoV-2 in China in December 2019, the world witnessed the emergence of the SARS-CoV-2 outbreak. Due to the high transmissibility rate of the virus, there is an urgent need to design and develop vaccines against SARS-CoV-2 to prevent more cases affected by the virus.

**Objective:** A computational approach is proposed for vaccine design against the SARS-CoV-2 spike (S) protein, as the key target for neutralizing antibodies, and envelope (E) protein, which contains a conserved sequence feature.

**Methods:** We used previously reported epitopes of S protein detected experimentally and further identified a collection of predicted B-cell and major histocompatibility (MHC) class II-restricted T-cell epitopes derived from E proteins with an identical match to SARS-CoV-2 E protein.

**Results:** The in silico design of our candidate vaccine against the S and E proteins of SARS-CoV-2 demonstrated a high affinity to MHC class II molecules and effective results in immune response simulations.

**Conclusions:** Based on the results of this study, the multiepitope vaccine designed against the S and E proteins of SARS-CoV-2 may be considered as a new, safe, and efficient approach to combatting the COVID-19 pandemic.

(*JMIR Bioinform Biotech* 2022;3(1):e36100) doi: [10.2196/36100](https://doi.org/10.2196/36100)

**KEYWORDS**

SARS-CoV-2; envelope protein; spike protein; COVID-19 vaccine; bioinformatics; COVID-19; informatics; immunoinformatics; computational model; vaccine design; pandemic

## Introduction

The recent outbreak of the new virus in Wuhan City, China, contributed to the discovery of a new coronavirus strain, labeled SARS-CoV-2, of the Coronaviridae family. This virus has caused severe damage and anxiety, leading to the loss of myriad individuals, impacting more than 535,863,950 people to date. SARS-CoV-2 causes the disease named COVID-19, which is associated with symptoms such as a flu-like illness, acute respiratory distress syndrome, and clinical or radiological evidence of pneumonia in individuals needing hospitalization [1]. Patients diagnosed with COVID-19 are reported to have high levels of interleukin (IL)1 $\beta$ , interferon (IFN) $\gamma$ , interferon-inducible protein 10 (IP10), and monocyte chemoattractant protein 1 (MCP1), likely leading to activated T helper-1 cell responses. In comparison, patients requiring intensive care unit admission had higher concentrations of granulocyte-colony stimulating factor, IP10, MCP1, MIP1A, and tumor necrosis factor- $\alpha$  than those not requiring intensive care, suggesting a possible correlation of cytokine storm and disease intensity. Nonetheless, SARS-CoV-2 infection also resulted in the enhanced production of T helper-2 cell cytokines such as IL4 and IL10, which inhibit inflammation that varies from that induced by SARS-CoV infection [2]. The persistent rise in patients and the high contagious rate of SARS-CoV-2 infection illustrate the immediate need to develop a safe and effective vaccine.

Vaccines are mostly comprised of whole pathogens, either destroyed or attenuated. However, it may be beneficial to use protein vaccines that are capable of generating an immune response against a specific pathogen. Epitope-based vaccines (EVs) utilize immunogenic proteins (epitopes) to induce an immune response. The performance of an EV is calculated by the number of epitopes to be used as the foundation. Nevertheless, the experimental identification of candidate epitopes is costly in terms of both time and money. Moreover, different immunological requirements need to be considered for the final choice of epitopes [3].

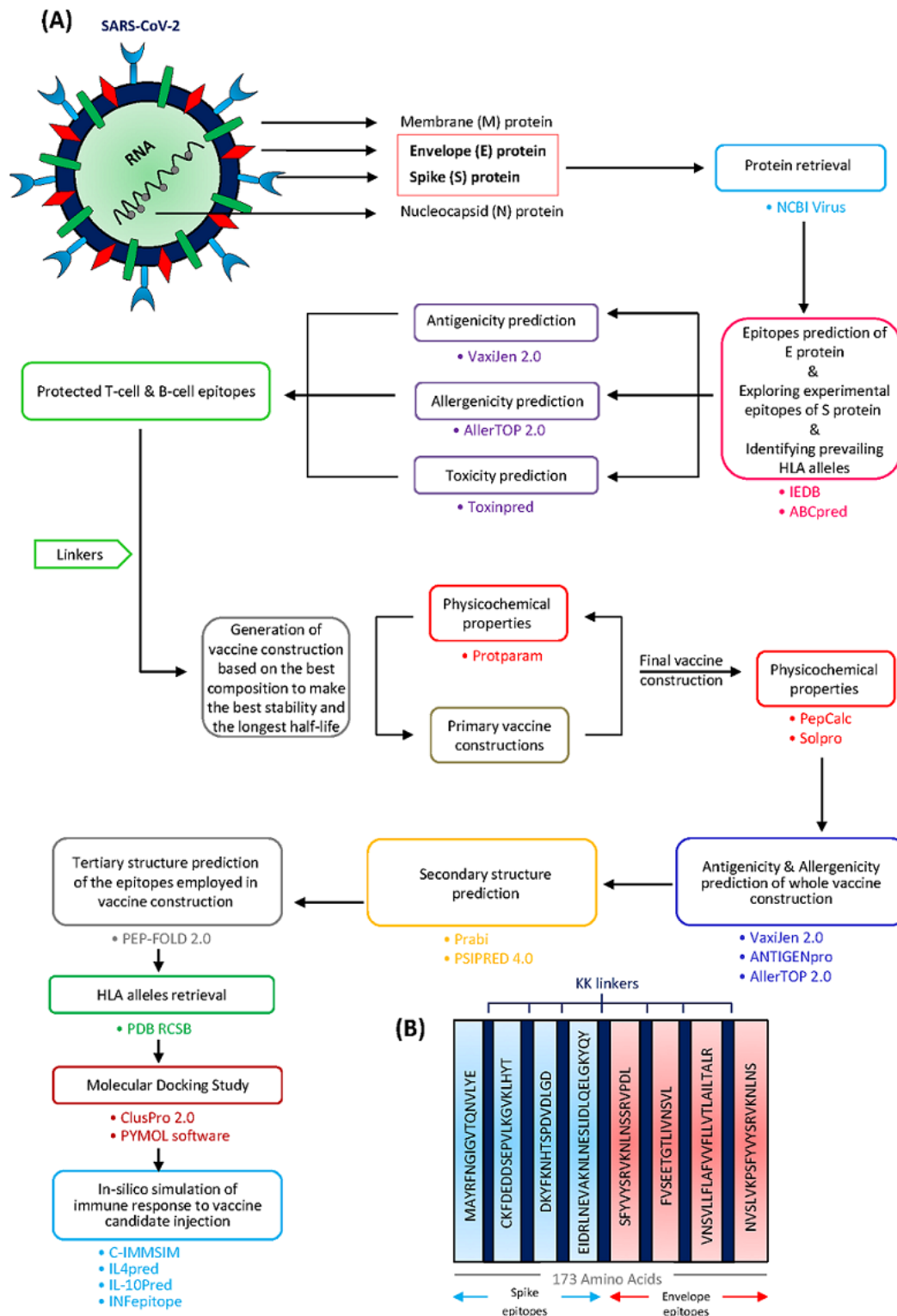
The properties of coronaviruses can be determined by electron microscopy. Coronaviruses are enveloped viruses with single-stranded positive-sense RNA. The coronavirus genome size varies from 26 to 32 kb [4]. Like all coronaviruses, SARS-CoV-2 comprises four viral proteins, namely spike (S) protein, a type of glycoprotein; membrane (M) protein, covering the membrane; envelope (E) protein, a strongly hydrophobic protein that covers the entire coronavirus structure; and

nucleocapsid (N) protein, a structural protein that suppresses RNA interference to overcome the host defense response [5,6] (Figure 1A). Such accessory proteins are not only essential for virion assembly but might also play additional roles in disrupting the host immune responses to promote viral replication [7]. SARS-CoV-2 requires the S glycoprotein, as the key target for neutralizing antibodies, to bind to the receptor and facilitate membrane fusion and virus entry. Every trimeric S protein monomer is roughly 180 kDa in size and comprises two subunits, S1 and S2, mediating binding and membrane fusion, respectively [8]. Therefore, S protein, but not other structural proteins, is the main antigen that causes the production of defensive neutralizing antibodies that stop viruses from attaching to their specific receptor, thereby preventing viral infection [9,10]. The S and M structural proteins have also been shown to have substantial mutational modifications, whereas the E and N proteins are highly conserved (Figure 1A), indicating differential selection pressures imposed on SARS-CoV-2 during evolution [11]. E protein is a small intrinsic membrane protein that is actively engaged in several stages of the life cycle of the virus, such as assembling, propagation, enveloping, and pathogenesis [12]. This protein also slows the transport of proteins through the secretive pathway by adjusting the concentrations of Ca<sup>2+</sup> and H<sup>+</sup> in the Golgi and endoplasmic reticulum compartments, which has been suggested as a mechanism for immune avoidance [13].

In this study, the S and E protein sequences were collected from a protein database and analyzed with various bioinformatics tools to identify protective epitopes. The toxicity of whole E protein as a second antigen was analyzed, and toxic epitopes were identified. The predicted B-cell and major histocompatibility complex (MHC) class II-restricted T-cell epitopes were checked in terms of not coinciding with these regions. The presence of less toxic epitopes in E protein in comparison with S protein served as a motivation to design an effective vaccine against these two antigens.

In particular, we sought to design a vaccine against the two structural antigens, S and E proteins, without using built-in adjuvants to obtain a vaccine that could be effective against all current and potential mutations of SARS-CoV-2, along with the advantage of a low molecular weight to avoid the complexity of future manufacturing. Since E protein is more highly conserved and the candidate vaccine showed all of the desired properties in simulations without using adjuvants, the study goals were achieved.

**Figure 1.** Schematic of the overall study design for development of a SARS-CoV-2 multi-epitope vaccine. (A) Study workflow of the in silico design of a multi-epitope vaccine against the envelope (E) protein of SARS-CoV-2. (B) Overlaps of 21 selected epitopes merged showing the final construct consisting of 8 epitopes. HLA: human leukocyte antigen; NCBI: National Center for Biotechnology Information.



## Methods

### Protein Sequence Retrieval

Based on the vaxquery database [14], the S and E proteins of SARS-CoV-2 were selected as targets for vaccine design because there are already vaccines in development or produced based on these proteins. The amino acid sequences of the E and

S proteins of SARS-CoV-2 were collected from the National Center for Biotechnology Information (NCBI) virus database [15] with accession number QHD43418 and QHR63280, respectively.

### B-Cell Epitopes Prediction

Prediction methods are both time- and cost-effective, and are reliable approaches for predicting linear B-cell epitopes as the

first step in the genome-wide quest for identifying B-cell antigens in a pathogenic organism [16]. The ABCpred database [17] includes the full-length E protein sequence for the prediction of linear B-cell epitopes. In this study, we set the “threshold” to 0.51 (default value), “length” to 16 (default value), and “overlapping filter” to “NO.” ABCpred uses a machine-learning methodology that requires fixed-length patterns for training or research, and the B-cell epitopes range from 5 to 30 residues in length. To overcome this issue, the server sought to create data sets of fixed-length patterns from B-cell epitopes by removing or linking residues to terminals. With a single hidden layer, the ABCpred server employs a partly recurrent neural network (Jordan network). The networks contain a single hidden layer with 35 residues and selectable window lengths of 10, 12, 14, 16, 18, and 20. The result is a single binary value that is either 1 or 0 (epitope or nonepitope).

The performance of prediction algorithms was evaluated using three parameters [17,18]: sensitivity, specificity, and accuracy. Sensitivity was calculated as the percentage of epitopes correctly identified as epitopes with the formula  $(TP/[TP+FN]) \times 100$ , where TP and FN are the numbers of true positives and false negatives, respectively. Specificity was calculated as the percentage of correctly predicted nonepitopes with the formula  $(TN/[TN+FP]) \times 100$ , where TN and FP are the numbers of true negatives and false positives, respectively. Accuracy is calculated as the total number of correct predictions (which includes both TP and TN) divided by the total number of forecasts made, multiplied by 100.

The Immune Epitope Database (IEDB) [19] was utilized to browse the available experimental B-cell assays on S protein of SARS-CoV-2. IEDB documents experimental evidence on antibody and T-cell epitopes examined in humans, nonhuman primates, and other animal species in the sense of infectious diseases, allergy, autoimmunity, and transplantation. We used the following parameters for browsing S epitopes: “Epitope” was set to “Linear peptide,” “Organism” was set to “SARS-CoV,” “Antigen” was set to “Spike glycoprotein,” “Assay” was set to “T cell” and “B Cell,” “MHC restriction” was set to “Class II,” and “Host” was set to “Human.”

For the B-cell epitope prediction of E protein using IEDB, we used the Bepiped Linear Epitope Prediction 2.0 service, which is based on a random forest algorithm trained on epitopes and nonepitope amino acids obtained from reported crystal structures to predict B-cell epitopes from a protein sequence, followed by sequential prediction smoothing [20].

### T-Cell Epitopes Prediction

T-cell epitopes are a group of proteins that can be detected by T-cell receptors after a given antigen has been processed intracellularly and attached to at least one MHC molecule, which are then expressed on the surface of antigen-presenting cells (APCs) as an MHC-protein complex. For entities that have at least one MHC molecule with strong affinity for binding to allergenic amino acid sequences from an allergen, the T-cell clones that can detect this MHC-protein complex are genetically susceptible to allergic reactions to this allergen. This concept can be investigated in silico by employing advanced statistical and mathematical methods [21]. The helper T lymphocyte (HTL)

epitopes of E protein were predicted by the IEDB database [19]. For T-cell MHC class II epitope prediction by IEDB, the “Prediction method” was set to “IEDB recommended 2.22,” “species/locus” was set to “human”/“HLA-DRA-DBR1\*01:01” and “HLA-DPB1\*01:02,” and “length” was left as the default setting. IEDB recommends using the consensus method, which compares a variety of methods to predict MHC class II epitopes, including a consensus approach combining NN-align, SMM-align, and combinatorial library methods [19]. The other tool that is available on the IEDB can browse the experimental HTL epitopes through the library based on a relevant antigen [19].

### Antigenicity, Allergenicity, and Toxicity Prediction

VaxiJen v2.0 with a threshold of 0.4 was used to predict the antigenicity of both B-cell and T-cell epitopes. VaxiJen is the first alignment-independent antigen predictor server, which was developed to achieve the categorization of antigens solely based on the physicochemical properties of proteins without recourse to sequence alignment. The system can be used either on its own or in conjunction with alignment-based prediction methods [22]. The methodology of this server is based on z descriptors, autocross covariance (ACC) preprocessing, discriminant analysis by partial least squares, and sequence similarity of the training set [23]. The z descriptors reflect the most critical physicochemical features for antigen recognition, including z1, z2, and z3 descriptors to describe the protein sequences. The hydrophobicity of amino acids is represented by the first principal component (z1), their size is represented by the second component (z2), and their polarity is represented by the third component (z3). The auto covariance  $A_{jj}(l)$  is represented by Equation (1) [22]:

$$A_{jj}(l) = \sum_i^{n-l} \frac{z_{j,i} \times z_{j,i+l}}{n-l} \quad (1)$$

The z-scales are calculated using index  $j$  ( $j=1, 2, 3$ ),  $n$  (number of amino acids in a sequence),  $I$  (amino acid position;  $I=1, 2, \dots, n$ ), and  $l$  (lag) ( $l=1, 2, \dots, L$ ). A small range of lags ( $L=1, 2, 3, 4, 5$ ) was employed to explore the effect of near amino acid proximity on protein antigenicity. Cross covariances  $C_{jk}(l)$  between two distinct z-scales,  $j$  and  $k$ , were calculated with Equation (2):

$$C_{jk}(l) = \sum_i^{n-l} \frac{z_{j,i} \times z_{k,i+l}}{n-l} \quad (2)$$

VaxiJen v2.0 was used to estimate the antigenicity of the whole-protein chimera. Based on this server, the antigenicity score of the final protein was 0.5830 (Probable ANTIGEN) with a threshold of 0.4. Likewise, ANTIGENpro [24] was utilized to predict the antigenicity of the protein chimera. ANTIGENpro is an alignment-free, sequence-based, and pathogen-independent protein antigenicity predictor. ANTIGENpro is the first indicator of protein antigenicity that is trained to employ reactivity data from the protein microarray analysis of five pathogens.



AllerTOP v2.0 was used to predict the allergenicity of both B-cell and HTL epitopes. Protein sequences are sent to this server in simple text. The results page then provides the identity of an allergen as “probable allergen” or “probable nonallergen.” The whole-protein chimera was predicted as a “probable nonallergen” using this tool [20]. This server was chosen because of its high sensitivity (94%) and higher rate of accurate prediction (94%-100%) in comparison to other similar servers to predict allergenicity [20]. Similar to VaxiJen, this database analyzes the presentation of protein sequences by z-descriptors and ACC transformation [17,18].

ToxinPred [25] with a protein fragment length of 10 was used to predict the toxicity of both B-cell and HTL epitopes. ToxinPred is a computational tool that was built to anticipate and design toxic versus nontoxic proteins. The primary data set used for this approach is comprised of 1805 toxic proteins ( $\leq 35$  residues). This server also was used to predict the toxicity of the whole-protein chimera and no fragment was predicted as a toxin.

### Construction of the Chimeric Protein

Selected B-cell and HTL epitopes were used to construct the protein chimera as a multiepitope vaccine. Overlaps of B-cell and HTL epitopes were merged. Bilysine (KK) linkers, as flexible linkers, were used to connect the epitopes. The KK linker was implanted between separate epitopes to maintain their independent immunological functions (Figure 1B). KK is the target sequence of cathepsin B, which is one of the essential antigen-processing proteases in MHC class II antigen presentation [26].

### Amino Acid Composition, Physicochemical Properties, and Solubility Prediction

The ProtParam database [27] was used to calculate and predict the molecular weight, isoelectric point (pI), in vivo and in vitro half-life, instability index II, and grand average of hydropathicity (GRAVY). ProtParam from the ExPASy server is a reliable algorithm to compute physicochemical properties. However, it uses a single sequence per analysis through the interface. The instability index is calculated using weight values, as shown in Equation (3) [28]:

$$II = (10/L) \sum_{i=1}^{L-1} DIWV(X_iY_{i+1} + 1)$$

(3)

where L is the length of the sequence and  $DIWV(x[i]x[i+1])$  is the instability weight value for the dipeptide starting in position  $i$ . A protein with an instability index less than 40 is anticipated to be stable, whereas one with an index greater than 40 is predicted to be unstable.

The relative volume occupied by aliphatic side chains (alanine, valine, isoleucine, and leucine) is known as the aliphatic index, which is considered a potentially beneficial element in the enhancement of globular protein thermostability. The aliphatic index is calculated by the formula  $X(\text{Ala})+aX(\text{Val})+b(X[\text{Ile}]+X[\text{Leu}])$  [29], where  $X(\text{Ala})$ ,  $X(\text{Val})$ ,  $X(\text{Ile})$ , and  $X(\text{Leu})$  represent the mole percent ( $100 \times$  mole fraction) of alanine, valine, isoleucine, and leucine, respectively,

and the coefficients  $a$  and  $b$  are the relative volumes of the valine side chain ( $a=2.9$ ) and Leu/Ile side chains ( $b=3.9$ ) to the side chain of alanine.

The SOLpro program from ANTIGENpro [24] was used to predict the solubility of protein chimera upon overexpression. SOLpro predicts the tendency of a protein to be soluble when overexpressed in *Escherichia coli* using a two-stage support vector machine model based on multiple representations of the primary sequence.

The PepCalc server [30] was used to predict the solubility of the final protein, which provides only a very rough estimation of water solubility.

### Secondary Structure Prediction

The Prabi server [31] was used to predict the secondary structure of the final sequence of the protein chimera. All PRABI components provide services in their various areas of expertise (eg, molecular, phylogeny, genomics, transcriptomics, proteomics, protein structure, and medical biostatistics). “GOR IV” was selected as the secondary structure prediction method. The program outputs two files: one with the sequence and anticipated secondary structure in rows (H=helix, E=extended or beta strand, and C=coil), and the other with the probability values for each secondary structure at each amino acid position (H=helix, E=extended or beta strand, and C=coil) [32].

The PSIPRED 4.0 [33] server was also used to predict the secondary structure, which provides more details of residues’ configurations. This is a very simple system of secondary prediction based on a simple neural network evaluation of PSI-BLAST-generated profiles, which is capable of generating findings that place the process at the very top of the prediction system crop [33].

### Molecular Docking of Final Vaccine Epitopes With MHC Molecules

PEP-FOLD 2.0 from the RPBS Web Portal server [32] was used to predict the tertiary structure of the vaccine construct epitopes. PEP-FOLD is an online tool that was designed to model 3D protein conformation structures in aqueous solutions for proteins 9-25 amino acids in length (de novo modeling). PEP-FOLD conducts a series of 50 simulations beginning with an amino acid sequence, and returns the most critical energy and population-related conformations found [32].

The ClusPro 2.0 server [34,35] rotates the ligands of each of the final epitopes of a vaccine protein with 70,000 rotations. The ligand rotations are translated relative to the MHC receptor alleles in three axes (x, y, z) on a grid. The top 1000 lowest energy docked structures from 70,000 rotations are then chosen and processed in turn. This set might have the potential to consist of at least some models that are close to the native structure of the complex. The server then clusters the 1000 rotations by finding the structure with the most “neighbors” within a 9 Å interface root mean square deviation radius as the distance measure. This ligand and its neighbors are then considered as the “cluster center” and the “members” of the cluster, respectively. This process was repeated for the remainder of the ligands to find the next clusters. Finally, the

server provides a score for the models and reports the top scoring models based on the cluster size (10 most populated clusters) [34,35]. One of the main advantages of ClusPro 2.0 as an automated protein docking server is its ability to generate protein-protein complexes with high accuracy [36].

PyMOL software was used to analyze the docking results. PyMOL is mostly utilized for molecular visualization by crystallographic, molecular dynamic simulation, and protein modeling software packages [37].

### Immune Response Simulation

IL4-, IL10-, and IFN $\gamma$ -inducing proteins from the 8 epitopes in the final vaccine construct were predicted via IL4pred server [38], IL-10Pred server [39], and IFNepitope server [40], respectively.

The immune response to vaccine injection was simulated using the C-ImmSim 10.1 server [41]. C-ImmSim 10.1 is an agent-based computational immune-response simulator that utilizes a position-specific score matrix and machine-learning methods for predicting epitope and immune interactions, respectively [42]. We regulated the parameters based on the predominant human leukocyte antigen (HLA) alleles of predictions. The host HLA selection parameters for MHC class I were set to A1010, A1101, and B0702; the parameters for DR MHC class II were set to DBR1\_0101; and the time step to injection was set to 1, 84, and 100 (maximum allowed value), respectively. We randomly shuffled the vaccine protein sequence (without adjuvants) using Stothard *P* 2000 from the Sequence Manipulation Suite server [43] to create a control group. The overall immunogenicity of the generic protein sequence associated with its amino acid sequence was assessed by this immune system simulation server [41]. The entire simulation was focused on three events, (1) B-cell epitopes binding, (2) HLA class I and II epitopes binding, and (3) T-cell receptor binding, in which the HLA-protein complex interaction should be present. Such processes are independently carried out by cells through various agents and the consumption of specific simulated biological quantities [41].

## Results

### Selection of Protein Sequences

The amino acid sequences of E protein and S protein of SARS-CoV-2 were collected from the NCBI virus database with accession numbers QHD43418 and QHR63280, respectively, which were released January 13, 2020, and have nucleotide completeness. The FASTA sequences were used to construct a multiepitope vaccine against SARS-CoV-2.

### B-Cell Epitopes Analysis

The ABCpred database reviewed the full-length E protein sequence for the analysis of linear B-cell epitopes. Among the results that passed the three filters of antigenicity, allergenicity, and toxicity, two epitopes (NVSLVKPSFYVYSRVK and YVYSRVKLNLSRVDPD) were chosen as protective epitopes (Table 1). The IEDB was then utilized to investigate the B-cell linear epitopes, resulting in 37 epitopes that were experimentally identified for S protein [44-54]. By contrast, there were no experimental B-cell epitopes for E protein of SARS-CoV-2.

### T-Cell Epitopes Analysis

The binding epitopes to MHC class II molecules of E protein were analyzed by the IEDB. We used the prediction method to identify T-cell epitopes of E protein since there were no corresponding experimental epitopes in this database, whereas we used the available experimental T-cell epitopes of S protein [55]. The same three filters of antigenicity, allergenicity, and toxicity were applied to identify the protective antigens. Based on the number of alleles, the predominant HLA alleles were HLA-DRA-DBR1\*01:01 and HLA-DPB1\*01:02 among MHC class II alleles.

### Antigenicity of Potential Epitopes

The antigenicity of both the B-cell and T-cell epitopes was predicted by VaxiJen 2.0, with a threshold of 0.4. The predicted epitopes with an antigenicity score above the threshold were considered as “antigen” epitopes. Screenings for the other two filters (allergenicity and toxicity) were not carried out on “nonantigen” epitopes (Table 1 and Table 2).

**Table 1.** Predicted T-cell and B-cell epitopes of SARS-CoV-2 envelope protein.<sup>a</sup>

Epitope	B-cell	MHC <sup>b</sup> II	Antigenicity	Allergenicity	Toxicity
<b>ABCpred</b>					
<b>Not selected for vaccine construction</b>					
TLAILTALRLCAYCCN	+ <sup>c</sup>	_ <sup>d</sup>	Antigen	Nonallergen	Toxin
LCAYCCNIVNVLVSKP	+	-	Antigen	Nonallergen	Toxin
FVSEETGTLIVNSVLL	+	-	Nonantigen	Discontinued	Discontinued
<b>Selected for vaccine construction</b>					
NVSLVKPSFYVYSRVK	+	-	Antigen	Nonallergen	Nontoxin
YVYSRVKLNLSRVPD	+	+	Antigen	Nonallergen	Nontoxin
<b>IEDB<sup>e</sup></b>					
<b>Not selected for vaccine construction</b>					
IVNSVLLFLAFVVFL	-	+	Antigen	Allergen	Discontinued
EETGTLIVNSVLLFL	-	+	Antigen	Allergen	Discontinued
GTLIVNSVLLFLAFV	-	+	Nonantigen	Discontinued	Discontinued
IVNSVLLFLAFVVFL	-	+	Nonantigen	Discontinued	Discontinued
MYSFVSEETGTLIVN	-	+	Nonantigen	Discontinued	Discontinued
NIVNVLVSKPSFYVY	-	+	Antigen	Allergen	Discontinued
RVKLNLSRVPDLLV	-	+	Antigen	Allergen	Discontinued
SEETGTLIVNSVLLF	-	+	Nonantigen	Discontinued	Discontinued
TGTLIVNSVLLFLAF	-	+	Nonantigen	Discontinued	Discontinued
YSFVSEETGTLIVNS	-	+	Nonantigen	Discontinued	Discontinued
<b>Selected for vaccine construction</b>					
SFYVYSRVKLNLSR	-	+	Antigen	Nonallergen	Nontoxin
FYVYSRVKLNLSRV	-	+	Antigen	Nonallergen	Nontoxin
YVYSRVKLNLSRVP	-	+	Antigen	Nonallergen	Nontoxin
FLAFVVFLVTLAIL	-	+	Antigen	Nonallergen	Nontoxin
FVSEETGTLIVNSVL	-	+	Antigen	Nonallergen	Nontoxin
KPSFYVYSRVKLNLS	-	+	Antigen	Nonallergen	Nontoxin
YSRVKLNLSRVPDL	-	+	Antigen	Nonallergen	Nontoxin
NSVLLFLAFVVFLV	-	+	Antigen	Nonallergen	Nontoxin
VKPSFYVYSRVKLNLS	-	+	Antigen	Nonallergen	Nontoxin
VNSVLLFLAFVVFL	-	+	Antigen	Nonallergen	Nontoxin
VSLVKPSFYVYSRVK	-	+	Antigen	Nonallergen	Nontoxin
VVFLVTLAILTALR	-	+	Antigen	Nonallergen	Nontoxin
LLFLAFVVFLVTLA	-	+	Antigen	Nonallergen	Nontoxin

<sup>a</sup>T-cell epitopes were identified as the best epitopes based on the number of alleles.

<sup>b</sup>MHC: major histocompatibility complex.

<sup>c</sup>+: Related.

<sup>d</sup>\_: Unrelated.

<sup>e</sup>IEDB: Immune Epitope Database.

**Table 2.** Experimental T-cell and B-cell epitopes of SARS-CoV-2 spike protein from Immune Epitope Database.

Epitope	Selected for vaccine construction	B-cell	MHC <sup>a</sup> II	Allergenicity	Toxicity
AATKMSECVLGQSKRVD	No	+ <sup>b</sup>	- <sup>c</sup>	Allergen	Discontinued
CKFDEDDSEPVKGVKLYHT	Yes	+	-	Nonallergen	Nontoxin
DDSEPVKGVKLYHT	Yes	+	-	Nonallergen	Nontoxin
DKYFKNHTSPDVLGD	Yes	+	-	Nonallergen	Nontoxin
DLGDISGINASVVNIQK	No	+	-	Allergen	Discontinued
EIDRLNEVAKNLNESLIDLQELGKYEYQ	Yes	+	-	Nonallergen	Nontoxin
EVAKNLNEIDLQELG	No	+	-	Allergen	Discontinued
KNHTSPDVLGDISGIN	No	+	-	Allergen	Discontinued
LYQDVNC	No	+	-	Allergen	Discontinued
LYQDVNCT	No	+	-	Allergen	Discontinued
MAYRFNGIGVTQNVLY	Yes	+	-	Nonallergen	Nontoxin
MAYRFNGIGVTQNVLYE	Yes	+	-	Nonallergen	Nontoxin
RASANLAATKMSECVLG	No	+	-	Allergen	Discontinued
SPDVLGDISGINAS	No	+	-	Allergen	Discontinued
MAYRFNGIGVTQNVLY	Yes	-	+	Nonallergen	Nontoxin
QLIRAAEIRASANLAATK	No	-	+	Allergen	Discontinued

<sup>a</sup>MHC: major histocompatibility complex.

<sup>b</sup>+: Related.

<sup>c</sup>-: Unrelated.

### Allergenicity of Potential Epitopes

The allergenicity of both B-cell and HTL epitopes was estimated by AllerTOP v. 2.0 (Table 1).

### Toxicity of Potential Epitopes

The toxicity of both B-cell and HTL epitopes was predicted by ToxinPred, with a protein fragment length of 10 (Tables 1 and 2).

### Construction of the Chimeric Protein

The screened epitopes were chosen for the design of a chimeric protein as a multiepitope vaccine. As shown in Tables 1 and 2, we selected 21 epitopes (including 6 B-cell epitopes and 1 HTL epitope of S protein, and 2 B-cell epitopes and 12 HTL epitopes of E protein), which all filled the criteria of “antigen,” “nonallergen,” and “nontoxin.” To establish a contiguous sequence in the final construction, the overlapping sequences of B-cell and T-cell epitopes were merged. In detail, MAYRFNGIGVTQNVLYE was obtained from MAYRFNGIGVTQNVLY (S protein HTL epitope), MAYRFNGIGVTQNVLY, and MAYRFNGIGVTQNVLYE (S protein B-cell epitopes); CKFDEDDSEPVKGVKLYHT was obtained from CKFDEDDSEPVKGVKLYHT and DDSEPVKGVKLYHT (S protein B-cell epitopes); SFYVYSRVKKNLNSRPDL was obtained from YVYSRVKKNLNSRPDL (E protein B-cell epitope), SFYVYSRVKKNLNSRP, YVYSRVKKNLNSRP, and YSRVKNLNSRPDL (E protein HTL epitopes); VNSVLLFLAFVVFLVTLAILTALR was obtained from

VVFLVTLAILTALR, LLFLAFVVFLVTLAILTA, FLAFVVFLVTLAIL, NSVLLFLAFVVFLV, and VNSVLLFLAFVVFLV (E protein HTL epitopes); and NVSLVKPSFYVYSRVKKNLNS was obtained from NVSLVKPSFYVYSRVK (E protein B-cell epitope), VSLVKPSFYVYSRVK, VKPSFYVYSRVKKNLNS, and KPSFYVYSRVKKNLNS (E protein HTL epitopes). Predicted linear B-cell epitopes and T-cell epitopes were connected utilizing KK linkers as flexible connectors (Figure 1B).

The arrangement of epitopes in the final vaccine construct had a substantial effect on the physicochemical properties such as half-life and instability, with the half-life varying from 5.5 hours to 30 hours in mammalian cells, and the stability varying from an unstable to a completely stable protein simply by changing the order of epitopes. Therefore, we further investigated the properties of more than 40 possible permutations considering the overlaps of the selected epitopes to find the best formulation of this vaccine candidate.

### Antigenicity, Allergenicity, and Toxicity Estimation of the Candidate Multiepitope Vaccine

The antigenicity of the final protein chimera (Figure 1B) was estimated by the VaxiJen 2.0 server to be 0.5830 with a threshold of 0.4. The ANTIGENpro platform was also utilized to estimate the antigenicity of the final protein. Based on this server, the whole protein (Figure 1B) is predicted as an antigen with a probability of 0.415508. The AllerTOP v.2.0 server indicated that the final protein is predicted as a “nonallergen”



and the ToxinPred server predicted the final protein as a “nontoxin.”

### Amino Acid Composition, Physicochemical Properties, and Solubility Prediction

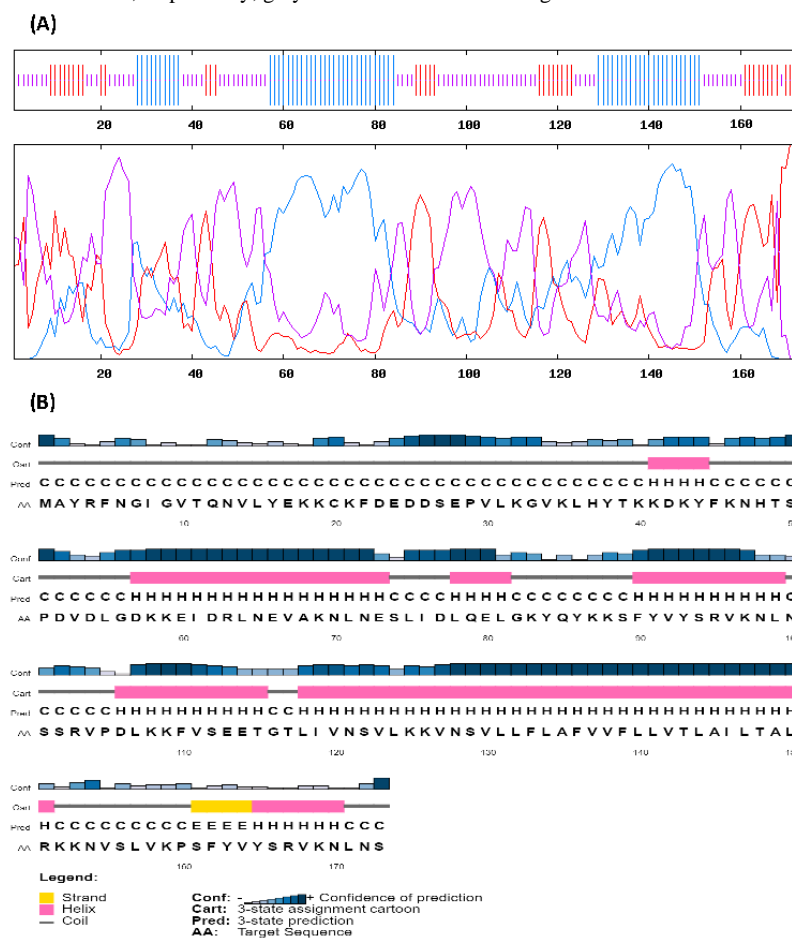
Based on the ProtParam database, the final protein chimera comprised 173 amino acids (Figure 1B) with a molecular weight of 19.9 kDa. The pI value was predicted to be 9.57. The half-life was estimated to be 30 hours in mammalian reticulocytes *in vitro*, more than 20 hours in yeast, and over 10 hours in *E. coli* *in vivo*. Instability index II was predicted to be 26.45, classifying the protein as stable (an index >40 indicates instability). The aliphatic index was 101.21, indicating high thermostability. The estimated GRAVY value was -0.293. This negative attribute

indicates that the protein is hydrophilic and can react with water molecules. Furthermore, based on the PepCalc server, the solubility was predicted to be “good” in water. Based on SOLpro from ANTIGENpro, the protein chimera was expected to be soluble with a probability of 0.765767.

### Secondary Structure Prediction

The secondary structure of the final protein (Figure 1B) was analyzed by the Prabi server. The final chimeric protein was estimated to include 35.26% alpha helices, 20.81% extended strands, and 43.93% random coils (Figure 2A). The detailed secondary structure predicted by the PSIPRED 4.0 server, including the residues and their configurations, is shown in Figure 2B.

**Figure 2.** Properties of the final vaccine protein candidate. (A) Secondary structure of the final protein predicted by the Prabi server. Red: extended strands; blue: alpha helix configurations; purple: random coils. (B) The residues and their arrangements analyzed by the PSIPRED 4.0 server. Yellow and pink regions indicate strands and helices, respectively; gray linkers indicate coil configurations.

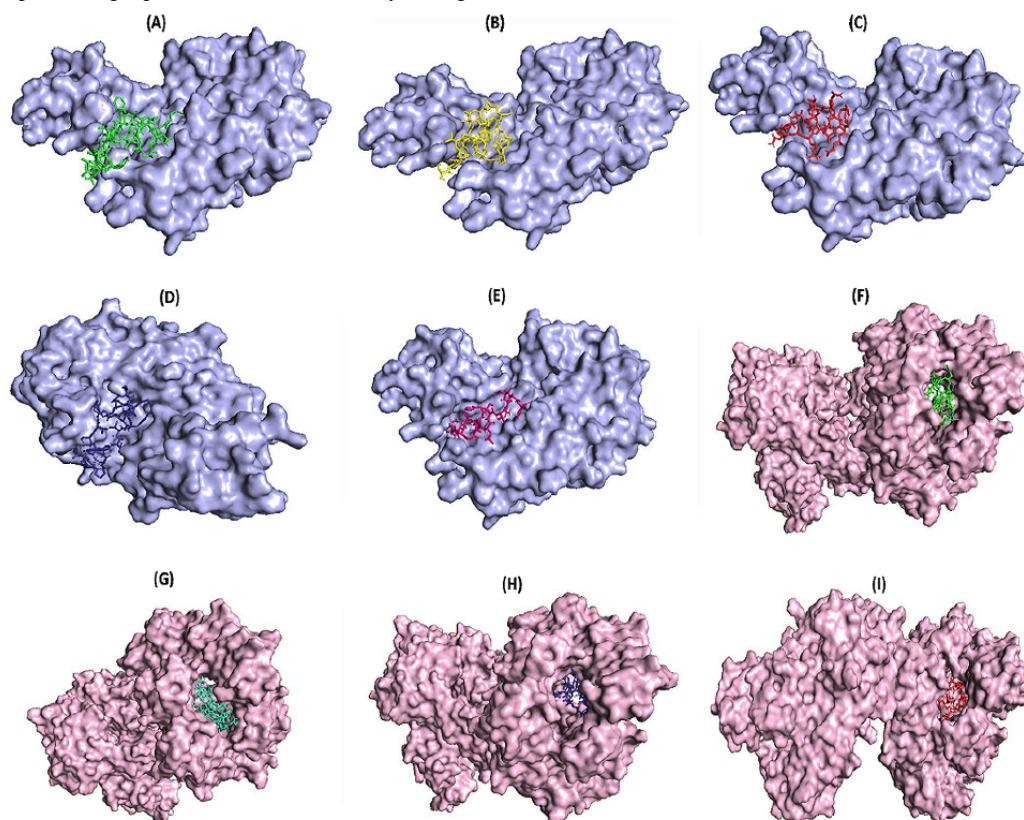


### Molecular Docking of Final Epitopes of the Vaccine With MHC Molecules

The crystal structures of HLA-DRA-DBR1\*01:01 and HLA-DPB1\*01:02 were retrieved from the PDB RCSB database (PDB ID: 1AQD and 3LQZ, respectively). The PDB files were edited and cleaned from heteroatoms. PEP-FOLD 2.0 from the RPBS web portal server was used to predict the tertiary structure of 8 epitopes of the vaccine construct individually. Molecular docking was performed on the epitopes and the whole vaccine construct with relevant MHC alleles using the ClusPro 2.0 online server. PyMOL software was used to perform a detailed analysis

of the interface of protein-protein interactions (Figure 3). The weighted score of the lowest energy docked complexes are reported in Table 3. The best way to rank the model is according to the cluster size (number of members) [34,35]. The most populated clusters were found in MAYRFNGIGVLTQNVLYE and HLA-DPA1\*01:03, DKYFKNHTSPVDLGD and HLA-DPA1\*01:03, CKFDEDDSEPV LKGVK LHYT and H L A - D P A 1 \* 0 1 : 0 3 , and EIDRLNEVAKNLNESLIDLQELGKYQY and HLA-DRB1\*01:01, with 784, 577, 350, and 261 cluster members, respectively.

**Figure 3.** Molecular docking analysis. (A) VNSVLLFLAFVVFLVTLAILTALR epitope (green) and HLA-DPA1\*01:03 protein (blue). (B) SFYVYSRVKLNSSSRVPDL epitope (yellow) and HLA-DPA1\*01:03 protein (blue). (C) MAYRFNGIGVTQNVLYE epitope (red) and HLA-DPA1\*01:03 protein (blue). (D) NVSLVKPSFYVYSRVKLNLS epitope (dark blue) and HLA-DPA1\*01:03 protein (blue). (E) FVSEETGTLIVNSVL epitope (pink) and HLA-DPA1\*01:03 protein (blue). (F) VNSVLLFLAFVVFLVTLAILTALR epitope (green) and HLA-DRB1\*01:01 protein (light pink). (G) EIDRLNEVAKNLSLIDLQELGKYQY epitope (light blue) and HLA-DRB1\*01:01 protein (light pink). (H) NVSLVKPSFYVYSRVKLNLS epitope (dark blue) and HLA-DRB1\*01:01 protein (light pink). (I) MAYRFNGIGVTQNVLYE epitope (red) and HLA-DRB1\*01:01 protein (light pink). HLA: human leukocyte antigen.



**Table 3.** Docking results and prediction of the immunity effects of epitopes.

Vaccine epitopes	Weighted scores <sup>a</sup> of the complex docked with				
	HLA <sup>b</sup> -DPA1*01:03	HLA-DRB1*01:01	IL <sup>c</sup> 4 inducer	IL10 inducer	IFN <sup>d</sup> $\gamma$ inducer
MAYRFNGIGVTQNVLYE	-869.2	-962.2	- <sup>e</sup>	-	-
CKFDEDDSEPVKGVKLYHT	-758.7	-810.6	+ <sup>f</sup>	-	+
DKYFKNHTSPDVLGD	-763.2	-755	+	-	-
EIDRLNEVAKNLSLIDLQELGKYQY	-715.4	-1050.7	-	+	+
SFYVYSRVKLNSSSRVPDL	-875.5	-992.2	+	+	+
FVSEETGTLIVNSVL	-798.2	-823.9	+	+	-
VNSVLLFLAFVVFLVTLAILTALR	-1148.3	-1477	-	+	+
NVSLVKPSFYVYSRVKLNLS	-852.7	-1036.6	+	+	+

<sup>a</sup>The weighted scores of the lowest energy docked structures were based on the cluster size of the most populated cluster.

<sup>b</sup>HLA: human leukocyte antigen.

<sup>c</sup>IL: interleukin.

<sup>d</sup>IFN: interferon.

<sup>e</sup>-: Unrelated.

<sup>f</sup>+: Related.

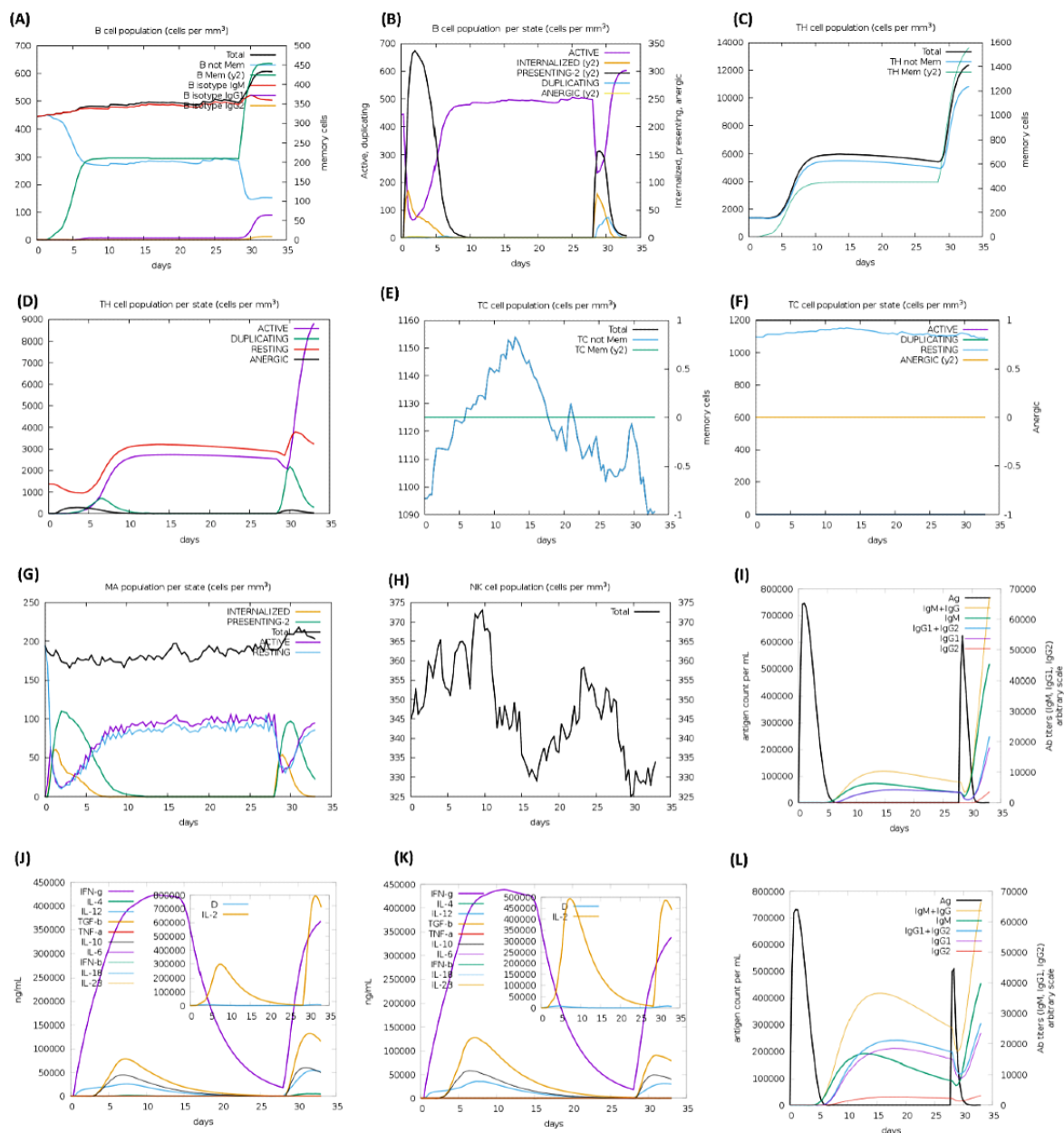
### Immune Response Simulation

We predicted the IL4, IL10, and IFN $\gamma$  inducing proteins from the 6 epitopes in the final vaccine construct via IL4pred server, IL10pred server, and IFNepitope server, respectively. The results are shown in Table 3.

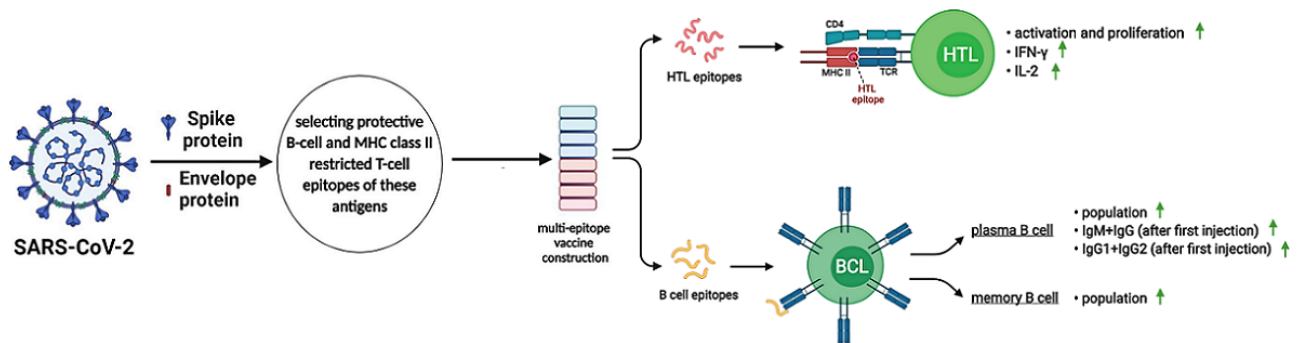
The primary and secondary immune responses were stimulated by the C-ImmSim 10.1 server. This server simulated the immune response of vaccine candidates with three injections in the time steps of 1, 84, and 100; each time step is equal to 8 hours. To

perform a relative comparison, we created a shuffled sequence of the vaccine candidate as a control protein, and we analyzed the results of the immune response simulation to the injection of the control. This shuffled sequence was employed to evaluate the significance of the vaccine sequence results, because in immune response simulation by this server, the sequence composition of the final epitopes connected via KK linkers is an important consideration. The results of the vaccine injection clearly varied from those of the controls (Figure 4 and Figure 5).

**Figure 4.** In silico immune response simulation to the injection of the candidate vaccine and control protein by the C-ImmSim 10.1 server. The simulation was performed with three injections in the time steps of 1, 84, and 100; each time step is equal to 8 hours. (A) B-cell population. (B) B-cell population per state. (C) T helper (TH) cell population. (D) TH cell population per state. (E) T cytotoxic (TC) cell population. (F) TC cell population per state. (G) Macrophage (MA) cell population. (H) Natural killer (NK) cell population. (I) Immunoglobulins. (J) Cytokines. (K) Cytokines after the protein control injection. (L) Immunoglobulins following the protein control injection.



**Figure 5.** Graphical representation of immune response simulation to the injection of the vaccine candidate against SARS-CoV-2 spike (S) and envelope (E) proteins. HTL: helper T lymphocyte; BCL: B-cell lymphoma; IgM: immunoglobulin M; IgG: immunoglobulin G; IFN- $\gamma$ : interferon gamma; IL-2: interleukin-2; TCR: T-cell receptor; MHC class II: major histocompatibility complex class II; CD4: cluster of differentiation 4. The image was created using the Biorender illustrator tool.



## Discussion

### Principal Findings

EVs offer a new strategy for the prophylactic and therapeutic use of pathogen-specific immunity [56]. A multiepitope vaccine consisting of a protein series or overlapping proteins has been proposed as an appropriate solution to the prevention and treatment of viral infections [57-62]. The perfect multiepitope vaccine should be engineered to include epitopes that can activate cytotoxic T lymphocytes, T-cells, and B-cells, and trigger successful responses to specific viruses [57].

We here present the *in silico* design of a potential multiepitope vaccine against the S and E proteins of SARS-CoV-2, which comprises both B-cell and HTL epitopes and can stimulate the immune system responses impressively. Immune interference is less likely to be a concern for multicomponent vaccines against a specific organism. For multitarget vaccinations, a strong response to one immune agent may reduce the otherwise marginal reaction to the second immunogen, and thus render the individual susceptible to infection with the pathogen corresponding to the second immune agent [63]. Since the SARS-CoV-2 S glycoprotein is surface-exposed and facilitates entry into host cells, it is the major priority of neutralizing antibodies against infection and the target of therapeutic and vaccine development [64,65]. S protein is also a primary focus for the design of subunit vaccines for SARS-CoV and Middle East Respiratory Syndrome (MERS)-CoV [66]. S trimers are widely coated with N-linked glycans, which are crucial for efficient folding and for modulating accessibility to host proteases and neutralizing antibodies [65-70]. E protein is conserved in all coronaviruses and covers the entire surface of SARS-CoV-2 (Figure 1A). There are fewer toxic epitopes of E protein than found for S protein. This finding was verified in the literature, in which E protein was explored in SARS-CoV in 2003 and, more recently, in MERS-CoV, demonstrating the retention of this protein in seven strains using the BioEdit Package tool and less toxic regions than in the S protein [12,57-62]. Several studies have examined the potential of coronaviruses with mutated E protein, focusing specifically on SARS- and MERS-CoV, as live attenuated vaccine candidates associated with hopeful results [12,71-75]. We obtained the FASTA sequence of the S and E proteins of SARS-CoV-2 from the NCBI database. B-cell and HTL epitopes of E protein were

predicted by different servers, whereas experimentally confirmed epitopes were utilized for S protein. The epitopes were screened based on the three filters of antigenicity, allergenicity, and toxicity. Therefore, we selected only protective epitopes. We merged the overlaps of B-cell and T-cell epitopes and fused them with appropriate flexible linkers. Previous studies reported that KK linkers preserve independent immune responses when they are inserted between epitopes [26] (Figure 1B).

The absence of allergenic properties of the proposed protein chimera further increases its potential as a vaccine candidate [76]. Finally, the whole-protein chimera was analyzed for antigenicity, allergenicity, and toxicity, which was predicted as an antigen [22], nonallergen [20], and nontoxin [25]. The pI was calculated to be 9.57, which shows that the final protein is alkaline. The vaccine protein construct was predicted as “soluble” upon expression in the *E. coli* host. The structural stability of a vaccine is known to be an essential aspect of its effectiveness, which can ensure the appropriate presentation of antigens and thus efficiently activate the immune system [77,78]. The instability index II of our candidate was calculated to be 26.45, which indicates that this protein is “stable.”

Secondary structure analysis predicted that the final protein consists of 35.26% alpha-helices, 20.81% extended strands, and 43.93% random coils. Essential types of “structural antigens” have been identified as natively unfolded protein regions and alpha-helical coil proteins. These two structural types, when examined in synthetic proteins, can fold into their native structure and are therefore recognized by antibodies naturally triggered in response to infection [76,79]. In the context of structural vaccinology, a molecular docking study was needed to predict the binding affinity of epitopes to the crystallized fragment of antibodies or MHC molecules [80,81]. To analyze the affinity of the final multiepitope vaccine to MHC molecules, we performed 16 docking simulations on the 8 epitopes of the final vaccine with MHC class II receptors. The results of docking analyses were notable, demonstrating the high affinity of the final epitopes of the vaccine construct to MHC molecules. The interface of protein-protein interactions was further considered using a visualization tool.

In the next step of designing a multiepitope vaccine, a systems vaccinology approach is beneficial in assessing the human complex immune response at different stages of biological structures [82]. Finally, we utilized an immune simulator server



to predict the primary and secondary responses of the immune system to three injections of the candidate vaccine. From the cytokines simulation plot, we noted an increase in the levels of IL-4 and IFN $\gamma$ , which is similar to the clinical features of COVID-19 patients reported by Huang et al [2] (Figure 4J). Appropriate activation of APCs, high production of memory cells due to the extensive activation of B-cells and T-cells, control and clearance of antigens due to the creation of cytokines by the participation of T helper memory cells, and the evident long-term memory persistence after three injections could confirm the efficiency of our candidate vaccine [83].

Finally, we selected one of the multiepitope vaccine candidates with the lowest molecular weight (shortest sequence length), which can potentially result in low-cost manufacturing and shorten production times [84].

### Comparison With Prior Work

Unlike most of the multiepitope vaccines that have been suggested during the COVID-19 pandemic, we preferred to design a vaccine without built-in adjuvants. Since adjuvants are necessary to increase the dosage efficacy by preventing the rapid degradation of proteins [85], tank-mixed adjuvants can be added to the final formulation. For instance, aluminum salts can be candidate adjuvants as they are used in various viral and bacterial vaccines and would be expected to enhance the antigen stability [86].

Adjuvants are effective in vaccine stability and can contribute to immunization reduction and enhanced antibody responses.

Although our suggested vaccine lacks a built-in adjuvant, it demonstrates the same stability and half-life estimation as previously reported candidates. Our vaccine construct also showed roughly the highest AI in comparison with other suggested multiepitope vaccines, along with the highest thermostability [29]. These benefits are only achieved by the careful arrangement of selected epitopes in designing this vaccine construct. This study thus demonstrates the importance of testing various permutations of epitopes in vaccine properties.

Finally, discharging our multiepitope vaccine from built-in adjuvants can demonstrate that the immune simulation for two injections (Figure 4) was induced because of the designed vaccine without interference from any nonspecific immunization against the adjuvants.

### Conclusion

The goal of this research was to suggest a computational method for predicting protective B-cell and T-cell epitopes of the E protein of SARS-CoV-2 accompanied by experimental epitopes of S protein to construct a chimeric protein vaccine candidate against this pandemic disease. The results demonstrated the high affinity of this chimeric protein to MHC molecules of the immune system, and the outputs of immune response simulation to the injection of this novel vaccine confirmed our findings. Thus, this multiepitope vaccine designed against the S and E proteins of SARS-CoV-2 utilizing immunoinformatics methods may be considered a new, safe, and efficient approach against SARS-CoV-2.

### Acknowledgments

The authors are grateful to Mrs Hajipour for her English language comments on the initial draft of the manuscript.

### Authors' Contributions

FG, FN, and VH developed the concept for the study. All authors discussed and designed the study. FG performed all parts of the study. RAC supervised the bioinformatics analyses. FN supervised the immunology experiments. HS, AHRSM, and MN contributed to acquisition and data analysis. VH supervised the whole project. FG wrote the first draft and all authors critically reviewed and approved the final version of the manuscript.

### Conflicts of Interest

None declared.

### References

1. Razai MS, Doerholt K, Ladhani S, Oakeshott P. Coronavirus disease 2019 (covid-19): a guide for UK GPs. *BMJ* 2020 Mar 05;368:m800. [doi: [10.1136/bmj.m800](https://doi.org/10.1136/bmj.m800)] [Medline: [32144127](https://pubmed.ncbi.nlm.nih.gov/32144127/)]
2. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* 2020 Feb 15;395(10223):497-506 [FREE Full text] [doi: [10.1016/S0140-6736\(20\)30183-5](https://doi.org/10.1016/S0140-6736(20)30183-5)] [Medline: [31986264](https://pubmed.ncbi.nlm.nih.gov/31986264/)]
3. Bing Z, Sakharkar K, Sakharkar M. In silico design of epitope-based vaccines. In: Dubitzky W, Wolkenhauer O, Cho KH, Yokota H, editors. *Encyclopedia of systems biology*. New York, NY: Springer; 2013:1003-1015.
4. Wu A, Peng Y, Huang B, Ding X, Wang X, Niu P, et al. Genome composition and divergence of the novel coronavirus (2019-nCoV) originating in China. *Cell Host Microbe* 2020 Mar 11;27(3):325-328 [FREE Full text] [doi: [10.1016/j.chom.2020.02.001](https://doi.org/10.1016/j.chom.2020.02.001)] [Medline: [32035028](https://pubmed.ncbi.nlm.nih.gov/32035028/)]
5. Pillaiyar T, Manickam M, Namasivayam V, Hayashi Y, Jung S. An overview of severe acute respiratory syndrome-coronavirus (SARS-CoV) 3CL protease inhibitors: peptidomimetics and small molecule chemotherapy. *J Med Chem* 2016 Jul 28;59(14):6595-6628 [FREE Full text] [doi: [10.1021/acs.jmedchem.5b01461](https://doi.org/10.1021/acs.jmedchem.5b01461)] [Medline: [26878082](https://pubmed.ncbi.nlm.nih.gov/26878082/)]



6. Cui L, Wang H, Ji Y, Yang J, Xu S, Huang X, et al. The nucleocapsid protein of coronaviruses acts as a viral suppressor of RNA silencing in mammalian cells. *J Virol* 2015 Sep;89(17):9029-9043 [FREE Full text] [doi: [10.1128/JVI.01331-15](https://doi.org/10.1128/JVI.01331-15)] [Medline: [26085159](https://pubmed.ncbi.nlm.nih.gov/26085159/)]
7. Pillaiyar T, Meenakshisundaram S, Manickam M. Recent discovery and development of inhibitors targeting coronaviruses. *Drug Discov Today* 2020 Apr;25(4):668-688 [FREE Full text] [doi: [10.1016/j.drudis.2020.01.015](https://doi.org/10.1016/j.drudis.2020.01.015)] [Medline: [32006468](https://pubmed.ncbi.nlm.nih.gov/32006468/)]
8. Ou X, Liu Y, Lei X, Li P, Mi D, Ren L, et al. Characterization of spike glycoprotein of SARS-CoV-2 on virus entry and its immune cross-reactivity with SARS-CoV. *Nat Commun* 2020 Mar 27;11(1):1620. [doi: [10.1038/s41467-020-15562-9](https://doi.org/10.1038/s41467-020-15562-9)] [Medline: [32221306](https://pubmed.ncbi.nlm.nih.gov/32221306/)]
9. Bisht H, Roberts A, Vogel L, Subbarao K, Moss B. Neutralizing antibody and protective immunity to SARS coronavirus infection of mice induced by a soluble recombinant polypeptide containing an N-terminal segment of the spike glycoprotein. *Virology* 2005 Apr 10;334(2):160-165 [FREE Full text] [doi: [10.1016/j.virol.2005.01.042](https://doi.org/10.1016/j.virol.2005.01.042)] [Medline: [15780866](https://pubmed.ncbi.nlm.nih.gov/15780866/)]
10. Bukreyev A, Lamirande EW, Buchholz UJ, Vogel LN, Elkins WR, St Claire M, et al. Mucosal immunisation of African green monkeys (*Cercopithecus aethiops*) with an attenuated parainfluenza virus expressing the SARS coronavirus spike protein for the prevention of SARS. *Lancet* 2004 Jun 26;363(9427):2122-2127 [FREE Full text] [doi: [10.1016/S0140-6736\(04\)16501-X](https://doi.org/10.1016/S0140-6736(04)16501-X)] [Medline: [15220033](https://pubmed.ncbi.nlm.nih.gov/15220033/)]
11. Ramaiah A, Arumugaswami V. Insights into cross-species evolution of novel human coronavirus 2019-nCoV and defining immune determinants for vaccine development. *bioRxiv*. 2020. URL: <https://www.biorxiv.org/content/10.1101/2020.01.29.925867v3> [accessed 2022-07-12]
12. Schoeman D, Fielding BC. Coronavirus envelope protein: current knowledge. *Viol J* 2019 May 27;16(1):69 [FREE Full text] [doi: [10.1186/s12985-019-1182-0](https://doi.org/10.1186/s12985-019-1182-0)] [Medline: [31133031](https://pubmed.ncbi.nlm.nih.gov/31133031/)]
13. de Jong AS, Visch H, de Mattia F, van Dommelen MM, Swarts HG, Luyten T, et al. The coxsackievirus 2B protein increases efflux of ions from the endoplasmic reticulum and Golgi, thereby inhibiting protein trafficking through the Golgi. *J Biol Chem* 2006 May 19;281(20):14144-14150 [FREE Full text] [doi: [10.1074/jbc.M511766200](https://doi.org/10.1074/jbc.M511766200)] [Medline: [16540472](https://pubmed.ncbi.nlm.nih.gov/16540472/)]
14. Xiang Z, Todd T, Ku KP, Kovacic BL, Larson CB, Chen F, et al. VIOLIN: vaccine investigation and online information network. *Nucleic Acids Res* 2008 Jan 23;36(Database issue):D923-D928 [FREE Full text] [doi: [10.1093/nar/gkm1039](https://doi.org/10.1093/nar/gkm1039)] [Medline: [18025042](https://pubmed.ncbi.nlm.nih.gov/18025042/)]
15. Brister J, Ako-Adjei D, Bao Y, Blinkova O. NCBI viral genomes resource. *Nucleic Acids Res* 2015 Jan;43(Database issue):D571-D577 [FREE Full text] [doi: [10.1093/nar/gku1207](https://doi.org/10.1093/nar/gku1207)] [Medline: [25428358](https://pubmed.ncbi.nlm.nih.gov/25428358/)]
16. Larsen J, Lund O, Nielsen M. Improved method for predicting linear B-cell epitopes. *Immunome Res* 2006 Apr 24;2:2 [FREE Full text] [doi: [10.1186/1745-7580-2-2](https://doi.org/10.1186/1745-7580-2-2)] [Medline: [16635264](https://pubmed.ncbi.nlm.nih.gov/16635264/)]
17. Saha S, Raghava GPS. Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. *Proteins* 2006 Oct 01;65(1):40-48. [doi: [10.1002/prot.21078](https://doi.org/10.1002/prot.21078)] [Medline: [16894596](https://pubmed.ncbi.nlm.nih.gov/16894596/)]
18. Saha S, Raghava G. Prediction methods for B-cell epitopes. In: Flower DR, editor. *Immunoinformatics. Methods in Molecular Biology*, vol 409. Totowa, NJ: Humana Press; 2007:387-394.
19. Vita R, Mahajan S, Overton J, Dhanda S, Martini S, Cantrell J, et al. The Immune Epitope Database (IEDB): 2018 update. *Nucleic Acids Res* 2019 Jan 08;47(D1):D339-D343 [FREE Full text] [doi: [10.1093/nar/gky1006](https://doi.org/10.1093/nar/gky1006)] [Medline: [30357391](https://pubmed.ncbi.nlm.nih.gov/30357391/)]
20. Dimitrov I, Bangov I, Flower DR, Doytchinova I. AllerTOP v.2--a server for in silico prediction of allergens. *J Mol Model* 2014 Jun 31;20(6):2278. [doi: [10.1007/s00894-014-2278-5](https://doi.org/10.1007/s00894-014-2278-5)] [Medline: [24878803](https://pubmed.ncbi.nlm.nih.gov/24878803/)]
21. Konstantinou G. T-Cell Epitope Prediction. *Methods Mol Biol* 2017;1592:211-222. [doi: [10.1007/978-1-4939-6925-8\\_17](https://doi.org/10.1007/978-1-4939-6925-8_17)] [Medline: [28315223](https://pubmed.ncbi.nlm.nih.gov/28315223/)]
22. Doytchinova IA, Flower DR. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinformatics* 2007 Jan 05;8:4 [FREE Full text] [doi: [10.1186/1471-2105-8-4](https://doi.org/10.1186/1471-2105-8-4)] [Medline: [17207271](https://pubmed.ncbi.nlm.nih.gov/17207271/)]
23. Wold S, Jonsson J, Sjöström M, Sandberg M, Rännar S. DNA and peptide sequences and chemical processes multivariately modelled by principal component analysis and partial least-squares projections to latent structures. *Analytica Chimica Acta* 1993 May;277(2):239-253. [doi: [10.1016/0003-2670\(93\)80437-p](https://doi.org/10.1016/0003-2670(93)80437-p)]
24. Cheng J, Randall AZ, Sweredoski MJ, Baldi P. SCRATCH: a protein structure and structural feature prediction server. *Nucleic Acids Res* 2005 Jul 01;33(Web Server issue):W72-W76 [FREE Full text] [doi: [10.1093/nar/gki396](https://doi.org/10.1093/nar/gki396)] [Medline: [15980571](https://pubmed.ncbi.nlm.nih.gov/15980571/)]
25. Gupta S, Kapoor P, Chaudhary K, Gautam A, Kumar R, Open Source Drug Discovery Consortium, et al. In silico approach for predicting toxicity of peptides and proteins. *PLoS One* 2013 Sep 13;8(9):e73957 [FREE Full text] [doi: [10.1371/journal.pone.0073957](https://doi.org/10.1371/journal.pone.0073957)] [Medline: [24058508](https://pubmed.ncbi.nlm.nih.gov/24058508/)]
26. Sarobe P, Lasarte J, Larrea E, Golvano J, Prieto I, Gullón A, et al. Enhancement of peptide immunogenicity by insertion of a cathepsin B cleavage site between determinants recognized by B and T cells. *Res Immunol* 1993 May;144(4):257-262. [doi: [10.1016/0923-2494\(93\)80102-5](https://doi.org/10.1016/0923-2494(93)80102-5)] [Medline: [7690980](https://pubmed.ncbi.nlm.nih.gov/7690980/)]
27. Gasteiger E, Hoogland C, Gattiker A, Wilkins M, Appel R, Bairoch A. Protein identification and analysis tools on the ExPASy server. In: Walker JM, editor. *The proteomics protocols handbook*. Totowa, NJ: Humana Press; 2005:571-607.
28. Guruprasad K, Reddy B, Pandit MW. Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence. *Protein Eng* 1990 Dec;4(2):155-161. [doi: [10.1093/protein/4.2.155](https://doi.org/10.1093/protein/4.2.155)] [Medline: [2075190](https://pubmed.ncbi.nlm.nih.gov/2075190/)]

29. Ikai A. Thermostability and aliphatic index of globular proteins. *J Biochem* 1980 Dec;88(6):1895-1898 [[FREE Full text](#)] [Medline: [7462208](#)]
30. Lear S, Cobb SL. Pep-Calc.com: a set of web utilities for the calculation of peptide and peptoid properties and automatic mass spectral peak assignment. *J Comput Aided Mol Des* 2016 Mar 24;30(3):271-277 [[FREE Full text](#)] [doi: [10.1007/s10822-016-9902-7](#)] [Medline: [26909892](#)]
31. Garnier J. GOR secondary structure prediction method version IV. *Meth Enzym* 1996;266:540-553. [doi: [10.1016/S0076-6879\(96\)66034-0](#)]
32. Maupetit J, Derreumaux P, Tuffery P. PEP-FOLD: an online resource for de novo peptide structure prediction. *Nucleic Acids Res* 2009 Jul 11;37(Web Server issue):W498-W503 [[FREE Full text](#)] [doi: [10.1093/nar/gkp323](#)] [Medline: [19433514](#)]
33. Jones DT. Protein secondary structure prediction based on position-specific scoring matrices. *J Mol Biol* 1999 Sep 17;292(2):195-202. [doi: [10.1006/jmbi.1999.3091](#)] [Medline: [10493868](#)]
34. Kozakov D, Beglov D, Bohnuud T, Mottarella SE, Xia B, Hall DR, et al. How good is automated protein docking? *Proteins* 2013 Dec;81(12):2159-2166 [[FREE Full text](#)] [doi: [10.1002/prot.24403](#)] [Medline: [23996272](#)]
35. Kozakov D, Hall DR, Xia B, Porter KA, Padhorny D, Yueh C, et al. The ClusPro web server for protein-protein docking. *Nat Protoc* 2017 Feb;12(2):255-278 [[FREE Full text](#)] [doi: [10.1038/nprot.2016.169](#)] [Medline: [28079879](#)]
36. Kozakov D, Hall DR, Beglov D, Brenke R, Comeau SR, Shen Y, et al. Achieving reliability and high accuracy in automated protein docking: ClusPro, PIPER, SDU, and stability analysis in CAPRI rounds 13-19. *Proteins* 2010 Nov 15;78(15):3124-3130 [[FREE Full text](#)] [doi: [10.1002/prot.22835](#)] [Medline: [20818657](#)]
37. Yuan S, Chan HS, Hu Z. Using PyMOL as a platform for computational drug design. *WIREs Comput Mol Sci* 2017 Jan 05;7(2):e1298. [doi: [10.1002/wcms.1298](#)]
38. Dhanda SK, Gupta S, Vir P, Raghava GPS. Prediction of IL4 inducing peptides. *Clin Dev Immunol* 2013;2013:263952. [doi: [10.1155/2013/263952](#)] [Medline: [24489573](#)]
39. Nagpal G, Usmani SS, Dhanda SK, Kaur H, Singh S, Sharma M, et al. Computer-aided designing of immunosuppressive peptides based on IL-10 inducing potential. *Sci Rep* 2017 Feb 17;7(1):42851. [doi: [10.1038/srep42851](#)] [Medline: [28211521](#)]
40. Dhanda SK, Vir P, Raghava GP. Designing of interferon-gamma inducing MHC class-II binders. *Biol Direct* 2013 Dec 05;8(1):30 [[FREE Full text](#)] [doi: [10.1186/1745-6150-8-30](#)] [Medline: [24304645](#)]
41. Rapin N, Lund O, Castiglione F. Immune system simulation online. *Bioinformatics* 2011 Jul 15;27(14):2013-2014. [doi: [10.1093/bioinformatics/btr335](#)] [Medline: [21685045](#)]
42. Nain Z, Abdulla F, Rahman M, Karim M, Khan MSA, Sayed SB, et al. Proteome-wide screening for designing a multi-epitope vaccine against emerging pathogen using immunoinformatic approaches. *J Biomol Struct Dyn* 2020 Oct;38(16):4850-4867. [doi: [10.1080/07391102.2019.1692072](#)] [Medline: [31709929](#)]
43. Stothard P. The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences. *Biotechniques* 2000 Jun;28(6):1102, 1104 [[FREE Full text](#)] [doi: [10.2144/00286ir01](#)] [Medline: [10868275](#)]
44. Poh CM, Carissimo G, Wang B, Amrun SN, Lee CY, Chee RS, et al. Two linear epitopes on the SARS-CoV-2 spike protein that elicit neutralising antibodies in COVID-19 patients. *Nat Commun* 2020 Jun 01;11(1):2806. [doi: [10.1038/s41467-020-16638-2](#)] [Medline: [32483236](#)]
45. Wang Q, Zhang L, Kuwahara K, Li L, Liu Z, Li T, et al. Immunodominant SARS coronavirus epitopes in humans elicited both enhancing and neutralizing effects on infection in non-human primates. *ACS Infect Dis* 2016 May 13;2(5):361-376 [[FREE Full text](#)] [doi: [10.1021/acsinfecdis.6b00006](#)] [Medline: [27627203](#)]
46. Buus S, Rockberg J, Forsström B, Nilsson P, Uhlen M, Schafer-Nielsen C. High-resolution mapping of linear antibody epitopes using ultrahigh-density peptide microarrays. *Mol Cell Proteomics* 2012 Dec;11(12):1790-1800 [[FREE Full text](#)] [doi: [10.1074/mcp.M112.020800](#)] [Medline: [22984286](#)]
47. Zhao J, Huang Q, Wang W, Zhang Y, Lv P, Gao X. Identification and characterization of dominant helper T-cell epitopes in the nucleocapsid protein of severe acute respiratory syndrome coronavirus. *J Virol* 2007 Jun;81(11):6079-6088 [[FREE Full text](#)] [doi: [10.1128/JVI.02568-06](#)] [Medline: [17392374](#)]
48. Lien S, Shih Y, Chen H, Tsai J, Leng C, Lin M, et al. Identification of synthetic vaccine candidates against SARS CoV infection. *Biochem Biophys Res Commun* 2007 Jul 06;358(3):716-721 [[FREE Full text](#)] [doi: [10.1016/j.bbrc.2007.04.164](#)] [Medline: [17506989](#)]
49. Chow SCS, Ho CYS, Tam TTY, Wu C, Cheung T, Chan PKS, et al. Specific epitopes of the structural and hypothetical proteins elicit variable humoral responses in SARS patients. *J Clin Pathol* 2006 May;59(5):468-476 [[FREE Full text](#)] [doi: [10.1136/jcp.2005.029868](#)] [Medline: [16461566](#)]
50. Hu H, Li L, Kao RY, Kou B, Wang Z, Zhang L, et al. Screening and identification of linear B-cell epitopes and entry-blocking peptide of severe acute respiratory syndrome (SARS)-associated coronavirus using synthetic overlapping peptide library. *J Comb Chem* 2005;7(5):648-656. [doi: [10.1021/cc0500607](#)] [Medline: [16153058](#)]
51. Chan WS, Wu C, Chow SCS, Cheung T, To K, Leung W, et al. Coronaviral hypothetical and structural proteins were found in the intestinal surface enterocytes and pneumocytes of severe acute respiratory syndrome (SARS). *Mod Pathol* 2005 Nov;18(11):1432-1439 [[FREE Full text](#)] [doi: [10.1038/modpathol.3800439](#)] [Medline: [15920543](#)]

52. Lai S, Chong PC, Yeh C, Liu LS, Jan J, Chi H, et al. Characterization of neutralizing monoclonal antibodies recognizing a 15-residues epitope on the spike protein HR2 region of severe acute respiratory syndrome coronavirus (SARS-CoV). *J Biomed Sci* 2005 Oct;12(5):711-727 [[FREE Full text](#)] [doi: [10.1007/s11373-005-9004-3](https://doi.org/10.1007/s11373-005-9004-3)] [Medline: [16132115](#)]
53. He Y, Zhou Y, Wu H, Luo B, Chen J, Li W, et al. Identification of immunodominant sites on the spike protein of severe acute respiratory syndrome (SARS) coronavirus: implication for developing SARS diagnostics and vaccines. *J Immunol* 2004 Sep 15;173(6):4050-4057 [[FREE Full text](#)] [doi: [10.4049/jimmunol.173.6.4050](https://doi.org/10.4049/jimmunol.173.6.4050)] [Medline: [15356154](#)]
54. Guo J, Petric M, Campbell W, McGeer PL. SARS corona virus peptides recognized by antibodies in the sera of convalescent cases. *Virology* 2004 Jul 01;324(2):251-256 [[FREE Full text](#)] [doi: [10.1016/j.virol.2004.04.017](https://doi.org/10.1016/j.virol.2004.04.017)] [Medline: [15207612](#)]
55. Yang J, James E, Roti M, Huston L, Gebe JA, Kwok WW. Searching immunodominant epitopes prior to epidemic: HLA class II-restricted SARS-CoV spike protein epitopes in unexposed individuals. *Int Immunol* 2009 Jan;21(1):63-71 [[FREE Full text](#)] [doi: [10.1093/intimm/dxn124](https://doi.org/10.1093/intimm/dxn124)] [Medline: [19050106](#)]
56. Khan AM, Miotto O, Heiny A, Salmon J, Srinivasan K, Nascimento EJ, et al. A systematic bioinformatics approach for selection of epitope-based vaccine targets. *Cellular Immunology* 2006 Dec;244(2):141-147. [doi: [10.1016/j.cellimm.2007.02.005](https://doi.org/10.1016/j.cellimm.2007.02.005)]
57. Zhang L. Multi-epitope vaccines: a promising strategy against tumors and viral infections. *Cell Mol Immunol* 2018 Feb 11;15(2):182-184 [[FREE Full text](#)] [doi: [10.1038/cmi.2017.92](https://doi.org/10.1038/cmi.2017.92)] [Medline: [28890542](#)]
58. Buonaguro L, HEPAVAC Consortium. Developments in cancer vaccines for hepatocellular carcinoma. *Cancer Immunol Immunother* 2016 Jan 21;65(1):93-99. [doi: [10.1007/s00262-015-1728-y](https://doi.org/10.1007/s00262-015-1728-y)] [Medline: [26093657](#)]
59. Brennick CA, George MM, Corwin WL, Srivastava PK, Ebrahimi-Nik H. Neoepitopes as cancer immunotherapy targets: key challenges and opportunities. *Immunotherapy* 2017 Mar;9(4):361-371 [[FREE Full text](#)] [doi: [10.2217/imt-2016-0146](https://doi.org/10.2217/imt-2016-0146)] [Medline: [28303769](#)]
60. Kuo T, Wang C, Badakhshan T, Chilukuri S, BenMohamed L. The challenges and opportunities for the development of a T-cell epitope-based herpes simplex vaccine. *Vaccine* 2014 Nov 28;32(50):6733-6745 [[FREE Full text](#)] [doi: [10.1016/j.vaccine.2014.10.002](https://doi.org/10.1016/j.vaccine.2014.10.002)] [Medline: [25446827](#)]
61. He R, Yang X, Liu C, Chen X, Wang L, Xiao M, et al. Efficient control of chronic LCMV infection by a CD4 T cell epitope-based heterologous prime-boost vaccination in a murine model. *Cell Mol Immunol* 2018 Sep 13;15(9):815-826 [[FREE Full text](#)] [doi: [10.1038/cmi.2017.3](https://doi.org/10.1038/cmi.2017.3)] [Medline: [28287115](#)]
62. Lu I, Farinelle S, Sausy A, Muller CP. Identification of a CD4 T-cell epitope in the hemagglutinin stalk domain of pandemic H1N1 influenza virus and its antigen-driven TCR usage signature in BALB/c mice. *Cell Mol Immunol* 2017 Jun 9;14(6):511-520 [[FREE Full text](#)] [doi: [10.1038/cmi.2016.20](https://doi.org/10.1038/cmi.2016.20)] [Medline: [27157498](#)]
63. Saul A, Fay MP. Human immunity and the design of multi-component, single target vaccines. *PLoS One* 2007 Sep 05;2(9):e850 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0000850](https://doi.org/10.1371/journal.pone.0000850)] [Medline: [17786221](#)]
64. Walls AC, Park Y, Tortorici MA, Wall A, McGuire AT, Veesler D. Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell* 2020 Apr 16;181(2):281-292 [[FREE Full text](#)] [doi: [10.1016/j.cell.2020.02.058](https://doi.org/10.1016/j.cell.2020.02.058)] [Medline: [32155444](#)]
65. Rossen JWA, de Beer R, Godeke G, Raamsman MJB, Horzinek MC, Vennema H, et al. The viral spike protein is not involved in the polarized sorting of coronaviruses in epithelial cells. *J Virol* 1998 Jan;72(1):497-503. [doi: [10.1128/jvi.72.1.497-503.1998](https://doi.org/10.1128/jvi.72.1.497-503.1998)]
66. Wang N, Shang J, Jiang S, Du L. Subunit vaccines against emerging pathogenic human coronaviruses. *Front Microbiol* 2020 Feb 28;11:298. [doi: [10.3389/fmicb.2020.00298](https://doi.org/10.3389/fmicb.2020.00298)] [Medline: [32265848](#)]
67. Walls AC, Tortorici MA, Frenz B, Snijder J, Li W, Rey FA, et al. Glycan shield and epitope masking of a coronavirus spike protein observed by cryo-electron microscopy. *Nat Struct Mol Biol* 2016 Oct 12;23(10):899-905 [[FREE Full text](#)] [doi: [10.1038/nsmb.3293](https://doi.org/10.1038/nsmb.3293)] [Medline: [27617430](#)]
68. Walls AC, Xiong X, Park Y, Tortorici MA, Snijder J, Quispe J, et al. Unexpected receptor functional mimicry elucidates activation of coronavirus fusion. *Cell* 2019 Feb 21;176(5):1026-1039 [[FREE Full text](#)] [doi: [10.1016/j.cell.2018.12.028](https://doi.org/10.1016/j.cell.2018.12.028)] [Medline: [30712865](#)]
69. Xiong X, Tortorici MA, Snijder J, Yoshioka C, Walls AC, Li W, et al. Glycan shield and fusion activation of a deltacoronavirus spike glycoprotein fine-tuned for enteric infections. *J Virol* 2018 Feb 15;92(4):e01628-17 [[FREE Full text](#)] [doi: [10.1128/JVI.01628-17](https://doi.org/10.1128/JVI.01628-17)] [Medline: [29093093](#)]
70. Yang Y, Liu C, Du L, Jiang S, Shi Z, Baric RS, et al. Two mutations were critical for bat-to-human transmission of Middle East Respiratory Syndrome coronavirus. *J Virol* 2015 Sep;89(17):9119-9123. [doi: [10.1128/jvi.01279-15](https://doi.org/10.1128/jvi.01279-15)]
71. Regla-Nava JA, Nieto-Torres JL, Jimenez-Guardeño JM, Fernandez-Delgado R, Fett C, Castaño-Rodríguez C, et al. Severe acute respiratory syndrome coronaviruses with mutations in the E protein are attenuated and promising vaccine candidates. *J Virol* 2015 Apr;89(7):3870-3887 [[FREE Full text](#)] [doi: [10.1128/JVI.03566-14](https://doi.org/10.1128/JVI.03566-14)] [Medline: [25609816](#)]
72. Netland J, DeDiego ML, Zhao J, Fett C, Álvarez E, Nieto-Torres JL, et al. Immunization with an attenuated severe acute respiratory syndrome coronavirus deleted in E protein protects against lethal respiratory disease. *Virology* 2010 Mar 30;399(1):120-128 [[FREE Full text](#)] [doi: [10.1016/j.virol.2010.01.004](https://doi.org/10.1016/j.virol.2010.01.004)] [Medline: [20110095](#)]
73. Almazán F, DeDiego ML, Sola I, Zuñiga S, Nieto-Torres JL, Marquez-Jurado S, et al. Engineering a replication-competent, propagation-defective Middle East respiratory syndrome coronavirus as a vaccine candidate. *mBio* 2013 Sep 10;4(5):e00650 [[FREE Full text](#)] [doi: [10.1128/mBio.00650-13](https://doi.org/10.1128/mBio.00650-13)] [Medline: [24023385](#)]

74. Lamirande EW, DeDiego ML, Roberts A, Jackson JP, Alvarez E, Sheahan T, et al. A live attenuated severe acute respiratory syndrome coronavirus is immunogenic and efficacious in golden Syrian hamsters. *J Virol* 2008 Aug;82(15):7721-7724 [FREE Full text] [doi: [10.1128/JVI.00304-08](https://doi.org/10.1128/JVI.00304-08)] [Medline: [18463152](https://pubmed.ncbi.nlm.nih.gov/18463152/)]
75. Fett C, DeDiego ML, Regla-Nava JA, Enjuanes L, Perlman S. Complete protection against severe acute respiratory syndrome coronavirus-mediated lethal respiratory disease in aged mice by immunization with a mouse-adapted virus lacking E protein. *J Virol* 2013 Jun 15;87(12):6551-6559. [doi: [10.1128/jvi.00087-13](https://doi.org/10.1128/jvi.00087-13)]
76. Shey RA, Ghogomu SM, Esoh KK, Nebangwa ND, Shintouo CM, Nongley NF, et al. In-silico design of a multi-epitope vaccine candidate against onchocerciasis and related filarial diseases. *Sci Rep* 2019 Mar 13;9(1):4409. [doi: [10.1038/s41598-019-40833-x](https://doi.org/10.1038/s41598-019-40833-x)] [Medline: [30867498](https://pubmed.ncbi.nlm.nih.gov/30867498/)]
77. Negahdaripour M, Nezafat N, Eslami M, Ghoshoon MB, Shoolian E, Najafipour S, et al. Structural vaccinology considerations for in silico designing of a multi-epitope vaccine. *Infect Genet Evol* 2018 Mar;58:96-109. [doi: [10.1016/j.meegid.2017.12.008](https://doi.org/10.1016/j.meegid.2017.12.008)] [Medline: [29253673](https://pubmed.ncbi.nlm.nih.gov/29253673/)]
78. Scheibelhofer S, Laimer J, Machado Y, Weiss R, Thalhamer J. Influence of protein fold stability on immunogenicity and its implications for vaccine design. *Expert Rev Vaccines* 2017 May;16(5):479-489 [FREE Full text] [doi: [10.1080/14760584.2017.1306441](https://doi.org/10.1080/14760584.2017.1306441)] [Medline: [28290225](https://pubmed.ncbi.nlm.nih.gov/28290225/)]
79. Bennuru S, Cotton JA, Ribeiro JMC, Grote A, Harsha B, Holroyd N, et al. Stage-specific transcriptome and proteome analyses of the filarial parasite *Onchocerca volvulus* and its endosymbiont. *mBio* 2016 Dec 30;7(6):e02028. [doi: [10.1128/mbio.02028-16](https://doi.org/10.1128/mbio.02028-16)]
80. Agostino M, Mancera RL, Ramsland PA, Fernández-Recio J. Optimization of protein-protein docking for predicting Fc-protein interactions. *J Mol Recognit* 2016 Nov;29(11):555-568. [doi: [10.1002/jmr.2555](https://doi.org/10.1002/jmr.2555)] [Medline: [27445195](https://pubmed.ncbi.nlm.nih.gov/27445195/)]
81. Ribas - Aparicio MR, Castelán - Vega JA, Jiménez - Alberto A, Monterrubio - López GP, Aparicio - Ozores G. The impact of bioinformatics on vaccine design and development. In: Afrin F, Hemeg H, Ozbak H, editors. *Vaccines*. London, UK: IntechOpen; 2017.
82. Raeven RHM, van Riet E, Meiring HD, Metz B, Kersten GFA. Systems vaccinology and big data in the vaccine development chain. *Immunology* 2019 Jan 13;156(1):33-46. [doi: [10.1111/imm.13012](https://doi.org/10.1111/imm.13012)] [Medline: [30317555](https://pubmed.ncbi.nlm.nih.gov/30317555/)]
83. Six A, Bellier B, Thomas-Vaslin V, Klatzmann D. Systems biology in vaccine design. *Microb Biotechnol* 2012 Mar;5(2):295-304. [doi: [10.1111/j.1751-7915.2011.00321.x](https://doi.org/10.1111/j.1751-7915.2011.00321.x)] [Medline: [22189033](https://pubmed.ncbi.nlm.nih.gov/22189033/)]
84. Strings DNA fragments. ThermoFisher Scientific. URL: <https://www.thermofisher.com/us/en/home/life-science/cloning/gene-synthesis/gene-strings-dna-fragments.html> [accessed 2022-07-12]
85. Sarkar I, Garg R, van Drunen Littel-van den Hurk S. Selection of adjuvants for vaccines targeting specific pathogens. *Expert Rev Vaccines* 2019 May 22;18(5):505-521 [FREE Full text] [doi: [10.1080/14760584.2019.1604231](https://doi.org/10.1080/14760584.2019.1604231)] [Medline: [31009255](https://pubmed.ncbi.nlm.nih.gov/31009255/)]
86. De Gregorio E, Caproni E, Ulmer JB. Vaccine adjuvants: mode of action. *Front Immunol* 2013;4:214. [doi: [10.3389/fimmu.2013.00214](https://doi.org/10.3389/fimmu.2013.00214)] [Medline: [23914187](https://pubmed.ncbi.nlm.nih.gov/23914187/)]

## Abbreviations

- ACC:** auto cross covariance
- APC:** antigen-presenting cell
- E protein:** envelope protein
- EV:** epitope-based vaccine
- FN:** false negative
- FP:** false positive
- GRAVY:** grand hydropathic average
- HLA:** human leukocyte antigen
- HTL:** helper T lymphocyte
- IEDB:** Immune Epitope Database
- IFN:** interferon
- IL:** interleukin
- IP10:** interferon-inducible protein 10
- KK:** bilysine
- MCP1:** monocyte chemoattractant protein 1
- MERS:** Middle East respiratory syndrome
- MHC:** major histocompatibility complex
- M protein:** membrane protein
- NCBI:** National Center for Biotechnology Information
- N protein:** nucleocapsid protein
- pI:** isoelectric point
- S protein:** spike protein
- TN:** true negative



**TP:** true positive

*Edited by A Mavragani; submitted 01.01.22; peer-reviewed by S Rostam Niakan Kalhori, A Banerjee; comments to author 27.04.22; revised version received 16.05.22; accepted 04.07.22; published 19.07.22*

*Please cite as:*

*Ghafouri F, Ahangari Cohan R, Samimi H, Hosseini Rad S M A, Naderi M, Noorbakhsh F, Haghpanah V*

*Development of a Multiepitope Vaccine Against SARS-CoV-2: Immunoinformatics Study*

*JMIR Bioinform Biotech 2022;3(1):e36100*

*URL: <https://bioinform.jmir.org/2022/1/e36100>*

*doi: [10.2196/36100](https://doi.org/10.2196/36100)*

*PMID:*

©Fateme Ghafouri, Reza Ahangari Cohan, Hilda Samimi, Ali Hosseini Rad S M, Mahmood Naderi, Farshid Noorbakhsh, Vahid Haghpanah. Originally published in JMIR Bioinformatics and Biotechnology (<https://bioinform.jmir.org>), 19.07.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Bioinformatics and Biotechnology, is properly cited. The complete bibliographic information, a link to the original publication on <https://bioinform.jmir.org/>, as well as this copyright and license information must be included.