

## Article

# Immunoinformatics-Based Identification of B and T Cell Epitopes in RNA-Dependent RNA Polymerase of SARS-CoV-2

Shabir Ahmad Mir <sup>1,\*</sup>, Mohammed Alaidarous <sup>1,2</sup>, Bader Alshehri <sup>1</sup>, Abdul Aziz Bin Dukhyil <sup>1</sup>, Saeed Banawas <sup>1,2,3</sup>, Yahya Madkhali <sup>1</sup>, Suliman A. Alsagaby <sup>1</sup> and Ayoub Al Othaim <sup>1</sup>

<sup>1</sup> Department of Medical Laboratory Sciences, College of Applied Medical Science, Majmaah University, Al Majmaah 11952, Saudi Arabia

<sup>2</sup> Health and Basic Sciences Research Center, Majmaah University, Al Majmaah 11952, Saudi Arabia

<sup>3</sup> Department of Biomedical Sciences, Oregon State University, Corvallis, OR 97331, USA

\* Correspondence: s.mir@mu.edu.sa; Tel.: +966-536300645

**Abstract:** Introduction: The ongoing coronavirus disease 2019 (COVID-19), which emerged in December 2019, is a serious health concern throughout the world. Despite massive COVID-19 vaccination on a global scale, there is a rising need to develop more effective vaccines and drugs to curb the spread of coronavirus. Methodology: In this study, we screened the amino acid sequence of the RNA-dependent RNA polymerase (RdRp) of SARS-CoV-2 (the causative agent of COVID-19) for the identification of B and T cell epitopes using various immunoinformatic tools. These identified potent B and T cell epitopes with high antigenicity scores were linked together to design the multi-epitope vaccine construct. The physicochemical properties, overall quality, and stability of the designed vaccine construct were confirmed by suitable bioinformatic tools. Results: After proper in silico prediction and screening, we identified 3 B cell, 18 CTL, and 10 HTL epitopes from the RdRp protein sequence. The screened epitopes were non-toxic, non-allergenic, and highly antigenic in nature as revealed by appropriate servers. Molecular docking revealed stable interactions of the designed multi-epitope vaccine with human TLR3. Moreover, in silico immune simulations showed a substantial immunogenic response of the designed vaccine. Conclusions: These findings suggest that our designed multi-epitope vaccine possessing intrinsic T cell and B cell epitopes with high antigenicity scores could be considered for the ongoing development of peptide-based novel vaccines against COVID-19. However, further in vitro and in vivo studies need to be performed to confirm our in silico observations.

**Keywords:** COVID-19; epitope; immunoinformatics; peptide vaccine; RNA polymerase; SARS-CoV-2



**Citation:** Mir, S.A.; Alaidarous, M.; Alshehri, B.; Bin Dukhyil, A.A.; Banawas, S.; Madkhali, Y.; Alsagaby, S.A.; Al Othaim, A. Immunoinformatics-Based Identification of B and T Cell Epitopes in RNA-Dependent RNA Polymerase of SARS-CoV-2. *Vaccines* **2022**, *10*, 1660. <https://doi.org/10.3390/vaccines10101660>

Academic Editor: Daniel Dory

Received: 7 September 2022

Accepted: 27 September 2022

Published: 3 October 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

In the last two decades, seven coronaviruses have infected human beings, with two main outbreaks caused by SARS-CoV (in 2002) and MERS-CoV (in 2012) [1]. In December 2019, the novel coronavirus (nCoV) emerged in Wuhan city in China, causing another global public health problem. This virus was first named as 2019-nCoV by WHO on January 12 and the associated disease was termed COVID-19 on 11 February 2020. Later, the virus causing the COVID-19 disease was identified as severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2), which is a member of the Coronaviridae family of viruses [2,3]. The world is currently experiencing severe circumstances of a worldwide public health emergency due to this viral pandemic. As of 28 August 2022, over 598 million confirmed cases of COVID-19, and over 6.4 million deaths by COVID-19, have been reported globally (WHO: <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports> (accessed on 28 August 2022)).

A lipid envelope containing the spike proteins and the membrane proteins surrounds the positive-stranded RNA genome of SARS-CoV-2. Although coronaviruses (CoVs) are

typically associated with respiratory illness, they can also cause infections in the brain and spinal cord [1,4]. The initial symptom of COVID-19 disease is fever, which can be accompanied with other symptoms such as shortness of breath, dry cough, muscle ache, headache, dizziness, sore throat, rhinorrhea, chest pain, nausea, diarrhea, and vomiting [5].

The SARS-CoV-2 binds to the host cell receptors through its spike proteins. The spike proteins also help the SARS-CoV-2 to release its viral genome into the host cell, where it is translated into two polyproteins and the structural proteins. Subsequently, the replication of the viral genome is started in the host cell [6]. Two-thirds of RNA encodes the two large non-structural polyproteins and the viral RNA-dependent RNA polymerase (RdRp) along with the associated accessory proteins. The remaining one-third of the viral genome codes for the four structural proteins (spike, membrane, envelope, and nucleocapsid) and the other helper proteins [7,8]. RdRp is derived from the polyproteins 1a and 1ab (encoded by ORF1a and ORF1ab, respectively) and is very important for the replication and transcription of the viral genome [9]. It is highly conserved among different RNA viruses. The core protein of RdRp consisting of a single chain of around 900 amino acid residues shows minimal activity. However, enhanced activity is achieved with the attachment of additional key subunits [9–11]. The development of effective treatment and vaccination strategies for SARS-CoV-2 is a high research priority worldwide. Various efforts are being carried out globally for the discovery of drugs and vaccines for the treatment of SARS-CoV-2 infection. Currently, 7 vaccines are approved by WHO for massive vaccination on a global scale, and over 300 additional vaccines are under development in the pre-clinical or clinical trial stage ([https://www.acrobiosystems.com/A1374-An-Overview-of-Different-COVID-19-Vaccines.html?gclid=EAlaIqobChMIn\\_dp5em-AIV2hXUAR2XOgJPEAAyAAEgKBCPD\\_BwE](https://www.acrobiosystems.com/A1374-An-Overview-of-Different-COVID-19-Vaccines.html?gclid=EAlaIqobChMIn_dp5em-AIV2hXUAR2XOgJPEAAyAAEgKBCPD_BwE) (accessed on 28 August 2022)).

The conventional methods for designing a vaccine (involving whole organisms or large proteins) lead to excessive antigenic load in addition to increasing the chances of allergic responses and toxicity [12]. This problematic concern can be overcome by designing peptide-based vaccines that encompass short immunogenic peptides with the ability to elicit strong and targeted immune responses, thereby avoiding the chances of allergenic and toxic reactions [12]. There are several recent studies available in the literature where the immunoinformatics-based vaccine designing efforts have been employed to address the challenge of vaccine construction for SARS-CoV-2 [13–25]. To add up to the available literature for developing an efficient peptide-based vaccine for SARS-CoV-2, here, we applied the *in silico* approach to identify the immunodominant B and T cell epitopes of the RdRp protein and develop a multi-epitope vaccine against SARS-CoV-2.

## 2. Methodology

### 2.1. Protein Sequence Retrieval and Determination of Antigenicity

The primary amino acid sequence of the RdRp protein of SARS-CoV-2 (accession number; YP\_009725307.1) was retrieved from the NCBI database (<https://www.ncbi.nlm.nih.gov/protein> (accessed on 25 October 2021)) and its antigenicity was checked by the ANTIGENpro tool of the Scratch protein predictor (<http://scratch.proteomics.ics.uci.edu/> (accessed on 25 October 2021)).

### 2.2. B Cell Epitope Prediction

We used the B cell epitope prediction server, BCPREDS server 1.1 (<http://ailab-projects1.ist.psu.edu:8080/bcpred/predict.html> (accessed on 10 November 2021)), with default settings of various parameters, to identify the linear B cell epitopes of the RdRp protein. The VaxiJen v2.0 server (<http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html> (accessed on 11 November 2021)) was used to evaluate the antigenicity of the predicted epitopes [26]. The epitopes with antigenicity scores above 0.4 were selected for further investigation.

### 2.3. Prediction of Helper T Lymphocyte (HTL) Epitopes and Their Screening

The IEDB server (<http://tools.iedb.org/mhcii/> (accessed on 15 November 2021)) was used for the prediction of MHCII binding/HTL epitopes in the protein sequence of RdRp [27]. Based on their percentile rank, the top 40 HTL epitopes were sorted. Epitopes with the lower percentile rank were considered to have higher MHCII binding affinity [28]. The selected epitopes were verified for their antigenicity by Vaxijen v2.0 [26] using the default threshold value of 0.4 for the prediction. The epitopes with an antigenicity score above 0.4 were screened for in silico IFN- $\gamma$  induction using the IFNepitope server.

The high-scoring epitopes (antigenicity score > 0.4) with the ability to induce IFN- $\gamma$  and cell-mediated immunity were identified by the IFNepitope server (<http://crdd.osdd.net/raghava/ifnepitope/> (accessed on 16 November 2021)). The prediction was made by a motif and support vector machine (SVM) hybrid approach [29]. Only the positive epitopes able to induce interferon- $\gamma$  (IFN- $\gamma$ ) were selected for further experiments.

### 2.4. Prediction of Cytotoxic T Lymphocyte (CTL) Epitopes

MHCI binding/CTL epitopes for supertypes A2, A3, and B7 were predicted in the RdRp sequence using an online NetCTL 1.2 server (<http://www.cbs.dtu.dk/services/NetCTL/> (accessed on 25 November 2021)) [30]. The epitopes were predicted by choosing default parameters for the three MHC I supertypes: A2, A3, and B7. The A2, A3, and B7 supertypes were selected because these three together cover >88% of the global inhabitants [27,30,31]. The default settings of NetCTL 1.2 (epitope identification threshold of 0.75, weight on TAP transport efficiency of 0.05, and weight on C terminal cleavage of 0.15) were chosen for epitope prediction. Here again, the VaxiJen v2.0 server was used to evaluate the antigenicity of the predicted epitopes. The CTL epitopes with an antigenicity score above 0.4 were selected for further investigation.

### 2.5. Toxicity Assessment and Prediction of the Allergenicity of the Screened Epitopes

The toxicity and allergenicity of all the screened HTL, CTL, and B cell epitopes were assessed by employing the ToxinPred server ([https://webs.iitd.edu.in/raghava/toxinpred/pep\\_test.php](https://webs.iitd.edu.in/raghava/toxinpred/pep_test.php) (accessed on 27 November 2021)) [27,32] and AllerTOP v2.0 server (<http://www.ddg-pharmfac.net/AllerTOP> (accessed on 28 November 2021)) [33], respectively. The default settings of both the servers were used for the respective predictions. The non-toxic and non-allergic HTL, CTL, and B cell epitopes with antigenicity scores more than 0.4 were selected for vaccine construction.

### 2.6. Construction of Multi-Epitope-Based Vaccine

The screened CTL, HTL, and B cell epitopes from the target protein were joined together through linkers (short and specific amino acid sequences) to generate the multi-epitope vaccine sequence. The different linkers used in this study include the B epitope linker (KK), HTL linker (GPGPG), and CTL linker (AAY).  $\beta$ -defensin 1 (Uniprot Id: P60022) was used as an adjuvant and linked through an EAAAK linker at the N-terminus of the vaccine sequence [34].

### 2.7. Prediction of Antigenicity and Allergenicity of the Designed Vaccine Candidate

The servers VaxiJen v2.0 (<http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html> (accessed on 10 December 2021)) and ANTIGENpro (<http://scratch.proteomics.ics.uci.edu/> (accessed on 11 December 2021)) were applied for the antigenicity evaluation of the designed multi-epitope vaccine construct [26,35]. The VaxiJen v2.0 server predicts the antigenicity on the basis of physicochemical properties of the input protein and the ANTIGENpro server predicts the antigenicity on the basis of protein microarray data analysis.

The server AllerTOP v2.0 (<http://www.ddg-pharmfac.net/AllerTOP> (accessed on 12 December 2021)) and AllergenFP (<https://ddg-pharmfac.net/AllergenFP/feedback.py> (accessed on 14 December 2021)) were used to predict the allergenicity of the designed multi-

epitopic vaccine [33,36]. AllerTOP v2.0 uses machine learning methods for the classification of allergens whereas AllergenFP uses SVM modules to predict the allergenicity of proteins.

### 2.8. Physicochemical Properties and Secondary Structure Prediction of the Constructed Vaccine

The ProtParam server (<http://web.expasy.org/protparam/> (accessed on 15 December 2021)) was used to predict the physicochemical properties of the designed vaccine construct [37] and PSIPRED server (<http://bioinf.cs.ucl.ac.uk/psipred/> (accessed on 16 December 2021)) [38] was used to predict its secondary structure [39].

### 2.9. Disorder Profile Generation

The disorder profile of the designed vaccine was produced through the DisProt server (<http://original.disprot.org/metapredictor.php> (accessed on 20 December 2021)) using the PONDR pool of four predictors [40–42] and IUPRED 2A predictor (<https://iupred2a.elte.hu/> (accessed on 20 December 2021)) [43–46].

### 2.10. Tertiary Structure Prediction and Validation

The tertiary structural model of the designed vaccine construct was generated using the I-TASSER web server (<https://zhanglab.ccmb.med.umich.edu/I-TASSER/> (accessed on 21 December 2021)) [47], which provides energy minimized models through iterative fragment assembly simulations, which are template-based. The topmost model was selected and evaluated for validation using ERRAT (<https://servicesn.mbi.ucla.edu/ERRAT/> (accessed on 30 December 2021)) and PROCHECK (<https://servicesn.mbi.ucla.edu/PROCHECK/> (accessed on 30 December 2021)) [48] web servers.

### 2.11. Molecular Interaction of the Designed Multi-Epitope Vaccine with TLR-3

The possibility of interaction between the immune receptor molecule (human TLR3) and the designed vaccine model was predicted by molecular docking experiments. The human TLR-3 protein structure (PDB ID: 1ZIW) was downloaded from the PDB database and the vaccine structure (predicted and evaluated models) was used as the ligand for this receptor. The interaction between the two was performed using the GRAMM-X simulation web server (<http://vakser.compbio.ku.edu/resources/gramm/grammx/> (accessed on 31 December 2021)) [49,50] and interactions were visualized using PyMOL software [<https://pymol.org/2/> (accessed on 2 January 2022)] [51].

### 2.12. In Silico Immune Stimulation Assay

To evaluate the immunological responses of the designed vaccine, its in silico immune simulation analysis was performed using C-ImmSim 10.1 server (<http://150.146.2.1/C-IMMSIM/index.php?page=0> (accessed on 5 January 2022)). The simulation was performed with default parameters, and we evaluated the immune stimulation for only one injection of the vaccine [52].

## 3. Results

### 3.1. B-Cell Epitopes

The RdRp protein is a probable antigen with a predicted antigenicity score of 0.620516 as revealed by the ANTIGENpro tool. The linear B cell epitopes of RdRp protein were predicted using BCPREDS server 1.1. The epitopes chosen on the basis of prediction scores were screened for their antigenicity using VaxiJen v2.0 server (Supplementary Table S1) and those with an antigen score of more than 0.4 were selected for further analysis. The BCPREDS server 1.1 analysis of the RdRp protein sequence revealed 13 B-cell epitopes among which only 5 epitopes possessed the antigenicity as revealed by the VaxiJen v2.0 server. Among these, three potential non-allergic, non-toxic, and non-overlapping epitopes with high antigenicity scores (>0.4) were selected for final vaccine construction (Table 1).

**Table 1.** Non-allergic and non-toxic antigenic B cell epitopes with respective antigenic scores.

S No	Starting Amino Acid Position	Peptide Epitope	Bcpreds Server Score	Antigenicity Score (On Vaxijen V2.0)
1	443	AQDGNAAISDYDYRY	0.987	0.5716
2	477	DKYFDCYDGGCINANQ	0.973	0.6443
3	347	HFRELGVVHNQDVNLH	0.938	1.4181

### 3.2. HTL Epitopes

The HTL epitopes of RdRp protein were predicted using the IEDB server for human MHC II alleles. The antigenic protein sequence of RdRp was analyzed on the T cell epitope-MHC II binding tool of IEDB while choosing the entire human HLA reference set and keeping other parameters of the server in the default setting. The rationale behind choosing the entire HLA reference set was that this would produce epitopes covering almost all the globally present MHC II alleles [27,28]. Based on their least percentile rank revealing their high affinity, the top 40 epitopes were selected and among these, 20 were antigenic as revealed by Vaxijen v2.0 server analysis (Supplementary Table S2). These 20 epitopes were further assessed for their IFN- $\gamma$ -inducing potential and only 15 epitopes were positive for IFN- $\gamma$  induction. All the 15 IFN- $\gamma$ -inducing epitopes were non-toxic, whereas one out of the 15 IFN- $\gamma$ -inducing epitopes proved to be a probable allergen. The final selection of epitopes was made on the basis of overlapping sequences and 10 epitopes with positive IFN- $\gamma$ -inducing potential and possessing high antigenicity scores (>0.4) were finally selected for vaccine construction (Table 2).

**Table 2.** Non-allergic and non-toxic antigenic HTL epitopes (with respective antigenic score and percentile rank) selected for constructing the multi-epitope vaccine. All these epitopes were positive for in silico IFN- $\gamma$  induction (run on the IFNepitope server).

S No.	Peptide Epitope	Starting Amino Acid Position	Percentile Rank	Antigenicity Score (On Vaxijen V2.0)
1	NMLRIMASLVLARKH	208	0.01	0.4897
2	PNMLRIMASLVLARK	207	0.01	0.4128
3	QMNLKYAISAKNRAR	121	0.07	1.5044
4	MNLKYAISAKNRART	122	0.08	1.4377
5	NLKYAISAKNRARTV	123	0.15	1.3422
6	DSYYSLLMPILTLTR	25	0.16	0.5134
7	MLRIMASLVLARKHT	209	0.22	0.5283
8	LRIMASLVLARKHTT	210	0.32	0.6646
9	LMPILTLTRALTAES	31	0.37	0.4225
10	ITQMNLKYAISAKNR	119	0.41	1.5061

### 3.3. CTL Epitopes

The CTL epitopes for supertypes A2, A3, and B7 were predicted using the NetCTL 1.2 server. A total of 35, 48, and 25 epitopes (for the supertypes A2, A3, and B7, respectively) were predicted in the RdRp protein sequence (Supplementary Table S3). Out of these, only 16, 28, and 13 were antigenic as revealed by the Vaxijen v2.0 server-based antigenicity analysis. All these epitopes were non-toxic in nature as analyzed by the ToxinPred server (Supplementary Table S4). Further selection of the epitopes was carried out based on their allergenicity. The server AllerTOP v2.0 ascertained only 18 CTL epitopes (5 for A2, 8 for A3, and 5 for B7) as probable non-allergens and hence good antigens for further experiments (Table 3).

**Table 3.** Non-allergic and non-toxic antigenic CTL epitopes (with respective antigenic scores) selected for designing the multi-epitope vaccine construct.

S No	Start	Peptide Epitope	Supertype	Antigenicity Score (On Vaxijen V2.0)
1	899	MLDMYSVML	A2	0.5626
2	467	RQLLFVVEV	A2	0.816
3	654	RLANECAQV	A2	0.9246
4	784	SVLYYQNNV	A2	0.5524
5	365	RLSFKELLV	A2	0.864
6	366	LSFKELLVY	A3	0.723
7	409	TVKPGNFNK	A3	1.378
8	10	RVCGVSAAR	A3	0.608
9	585	ATVVIGTSK	A3	0.757
10	543	NLKYAISAK	A3	1.351
11	365	RLSFKELLV	A3	0.864
12	341	VVSTGYHFR	A3	1.474
13	890	KLHDELTGH	A3	0.4147
14	581	ATRGATVVI	B7	0.569
15	242	MPILTLTRA	B7	0.806
16	579	IAATRGATV	B7	0.8883
17	399	AALTNNVAF	B7	0.638
18	528	FAYTKRNV	B7	1.0296

### 3.4. Multi-Epitope Vaccine Design

The selected antigenic peptides/epitopes were joined together through the linkers to create the multi-epitope vaccine candidate. In addition to maintaining the construct immunogenicity, the linkers also make the resulting protein complex more stable and flexible [27,28]. Our designed vaccine construct comprised 18 CTL epitopes, 10 HTL epitopes, and 3 B cell epitopes (Supplementary Figure S1). Additionally, for improving the immunogenicity of the multi-epitope vaccine, we tagged the construct with an adjuvant ( $\beta$ -defensin) at the N terminus.

### 3.5. Prediction of the Antigenicity and Allergenicity of the Constructed Vaccine

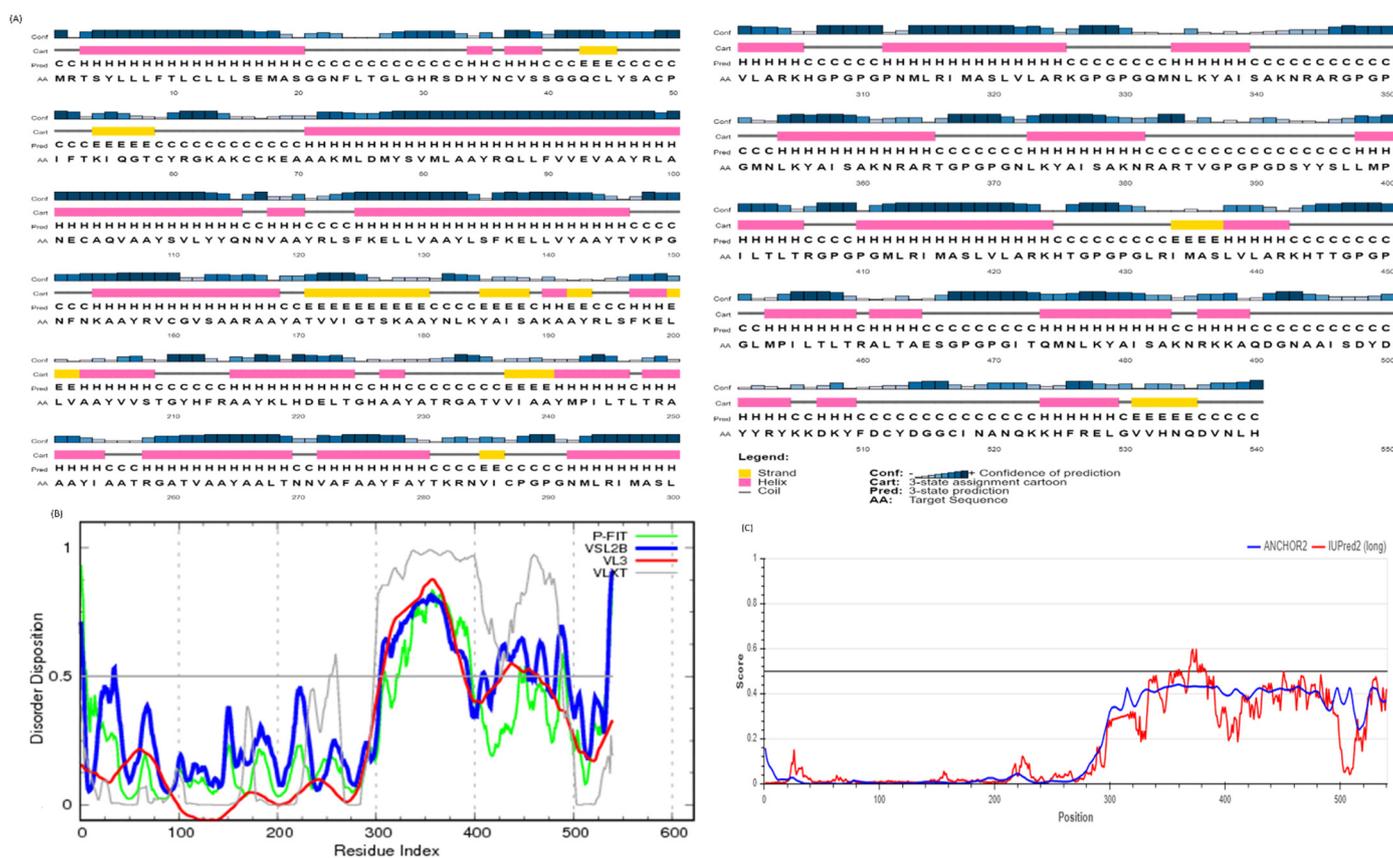
One of the characteristics of an ideal vaccine is that the vaccine must be antigenic and non-allergic in nature. Our designed vaccine was observed to be antigenic with antigenicity score of 0.7006 as predicted by the VaxiJen v2.0 server. The construct was also identified as a probable non-allergen by the AllerTOP v2.0 server.

### 3.6. Physicochemical Characterization of the Designed Vaccine

As revealed by ProtParam server analysis, the designed vaccine construct with 540 amino acids has a predicted molecular weight of 58.245 kDa. It is basic in nature (theoretical pI 9.93) and would possess a half-life of 30 h in mammalian reticulocytes (in vitro) and >10 h in *Escherichia coli* (in vivo). Its predicted instability index of 21.73 classifies the protein as stable. The designed vaccine is thermostable as revealed by its aliphatic index (88.17) and its average hydrophobicity is 0.034.

### 3.7. Prediction of Secondary Structure and Disordered Residues

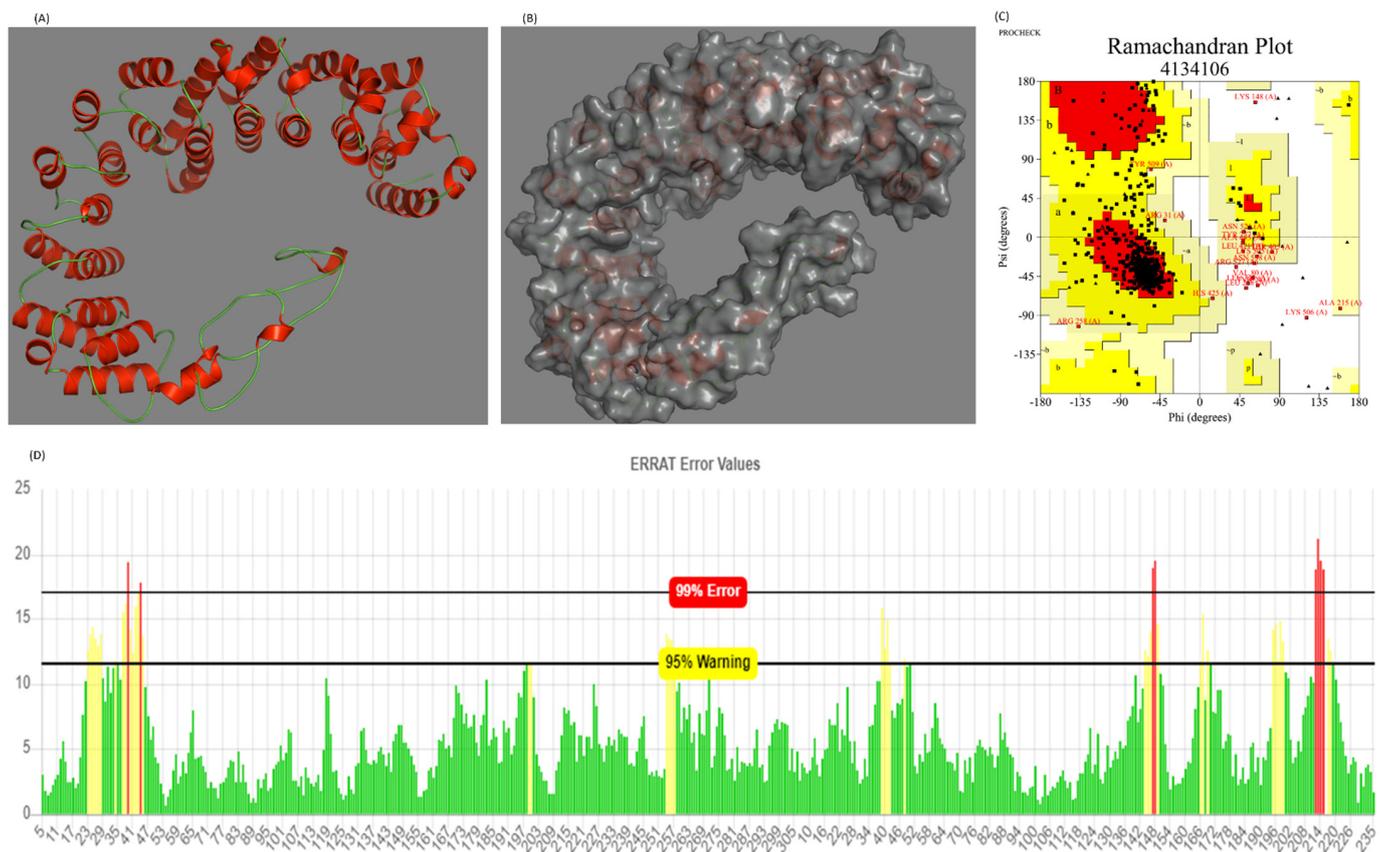
The secondary structure of the designed vaccine construct as predicted by the PEPISPRED 4.0 server contains several  $\alpha$ -helices, few  $\beta$ -strands, and many coils (Figure 1). Its disorder profiles, as generated through the DisProt server and IUPRED 2A predictor, are also shown in Figure 1.



**Figure 1.** (A) Secondary structure prediction of the designed vaccine construct (graphical result) as revealed by the PSIPRED server. The secondary structures are indicated by different colors:  $\alpha$ -helix (shown in pink),  $\beta$ -sheet (shown in yellow), and coil (shown in grey). (B) Intrinsic disorder in residues of the designed vaccine construct as predicted by PONDR® pool of predictors (PONDR-FIT, PONDR-VSL2B, PONDR-VL3, PONDR-VLXT). (C) Disorder in residues of the designed vaccine construct as predicted by the IUPRED 2A server.

### 3.8. Prediction and Quality Assurance of the Tertiary Structure of the Designed Vaccine

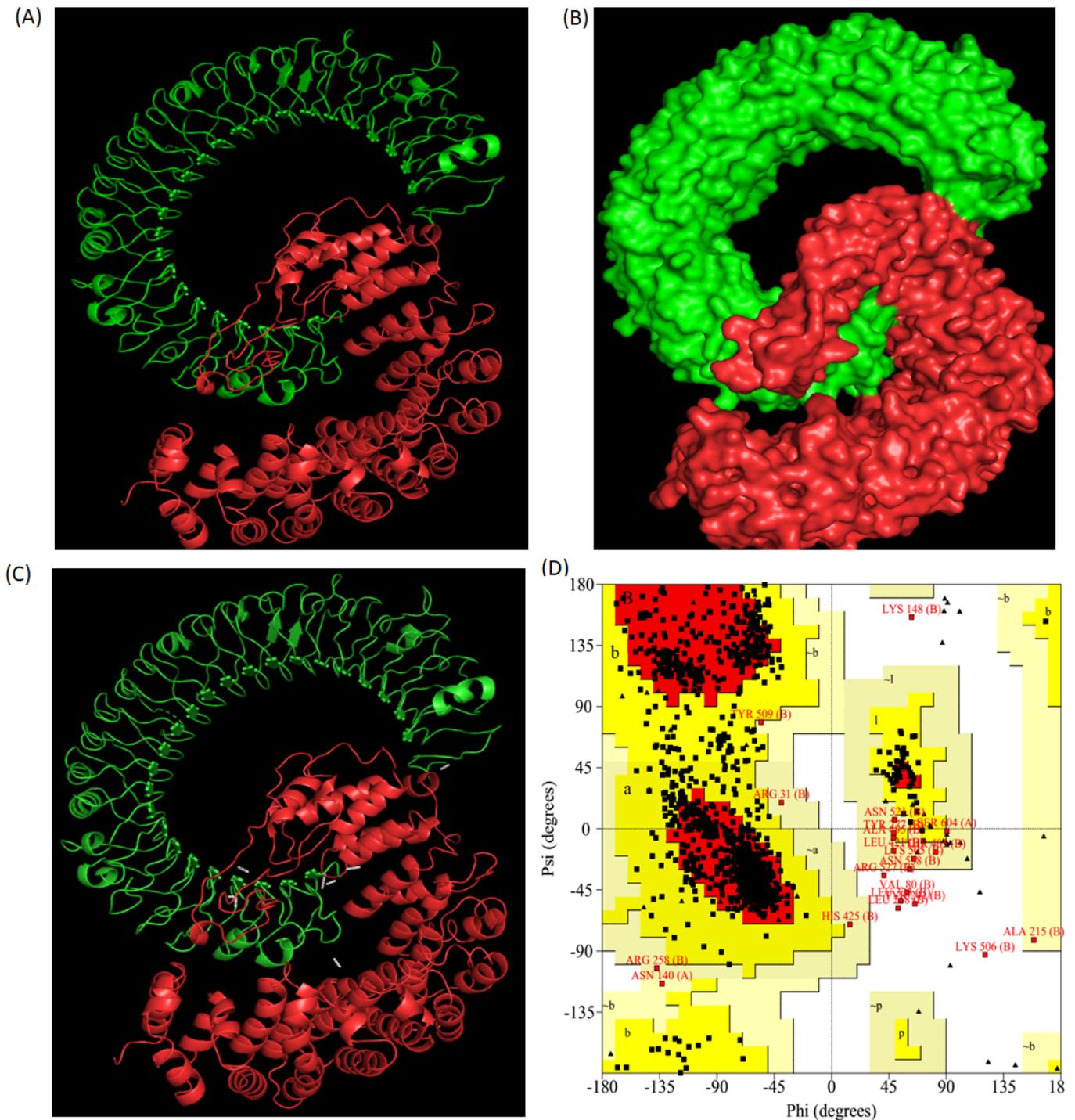
The designed 540-residue-long vaccine construct was modeled on the I-TASSER web server. The server revealed the five top models with C-scores ranging from  $-1.01$  to  $-3.77$ , among which the topmost model with the highest C-score ( $-1.01$ ) was selected for further analysis. A C-score with a higher value indicates a model with a greater confidence and vice versa. The selected model majorly comprised alpha helices and loops, with almost no contribution from beta sheets (Figure 2). To validate the structural quality of the predicted model, PROCHECK and ERRAT analyses were performed. The PROCHECK generated a Ramachandran plot of the 3D model of the vaccine, showing that 98.5% of the residues were localized in the favored regions and only 1.5% of them were positioned in the disallowed regions, which confirms the overall quality of the vaccine model (Figure 2). Ideally, a reliable model has at least 90% of its residues in the favored region [53]. The overall quality factor of the designed vaccine model as revealed by ERRAT analysis was 91.149, which signifies a good-quality model worth investigating further (Figure 2). An ERRAT score greater than 50 denotes a good-quality model [54].



**Figure 2.** (A,B) Structure of the selected 3D model of the final vaccine construct (cartoon and surface representation, respectively). Red and green represent the helical and loop region, respectively. (C) Ramachandran plot of the predicted model of the designed vaccine using the PROCHECK server (76.3%, 19.6%, 2.6%, and 1.5 in the favored, allowed, outlier, and disallowed regions, respectively). (D) Validation of the overall quality of the predicted model on ERRAT2 analysis (overall quality factor = 91.149).

### 3.9. Molecular Docking of the Vaccine Construct with TLR-3

An appropriate interaction between the antigen molecules and the immune receptor is essential for generating an immune response. Therefore, to ensure that the designed vaccine construct is capable of interacting with the immune receptor TLR-3, we executed a docking study using the GRAMM-X simulation web server. This docking server is exclusively designed for docking pairs of protein molecules. The protein-protein docking model complex produced by the GRAMM-X simulation web server and visualized by PyMOL is shown in Figure 3. The interactions between TLR-3 protein (chain A) and the designed vaccine construct (chain B) were also visualized by PyMOL software (see Figure 3). The PROCHECK-generated Ramachandran plot analysis of the protein-protein docking model complex showed that 99.3% of the residues were localized in the favored regions whereas only 0.7% of them were positioned in the disallowed regions, which verifies the overall quality of the docking complex (Figure 3). Further, the molecular dynamics simulations of the docked complex were performed on the CABS-flex server (<http://biocomp.chem.uw.edu.pl/CABSflex/> (accessed on 8 January 2022)). The root-mean-square fluctuation (RMSF) of the amino acid residues of both the chains of the complex (chain A, TLR3; chain B, designed vaccine construct) are shown in Supplementary Figure S2.

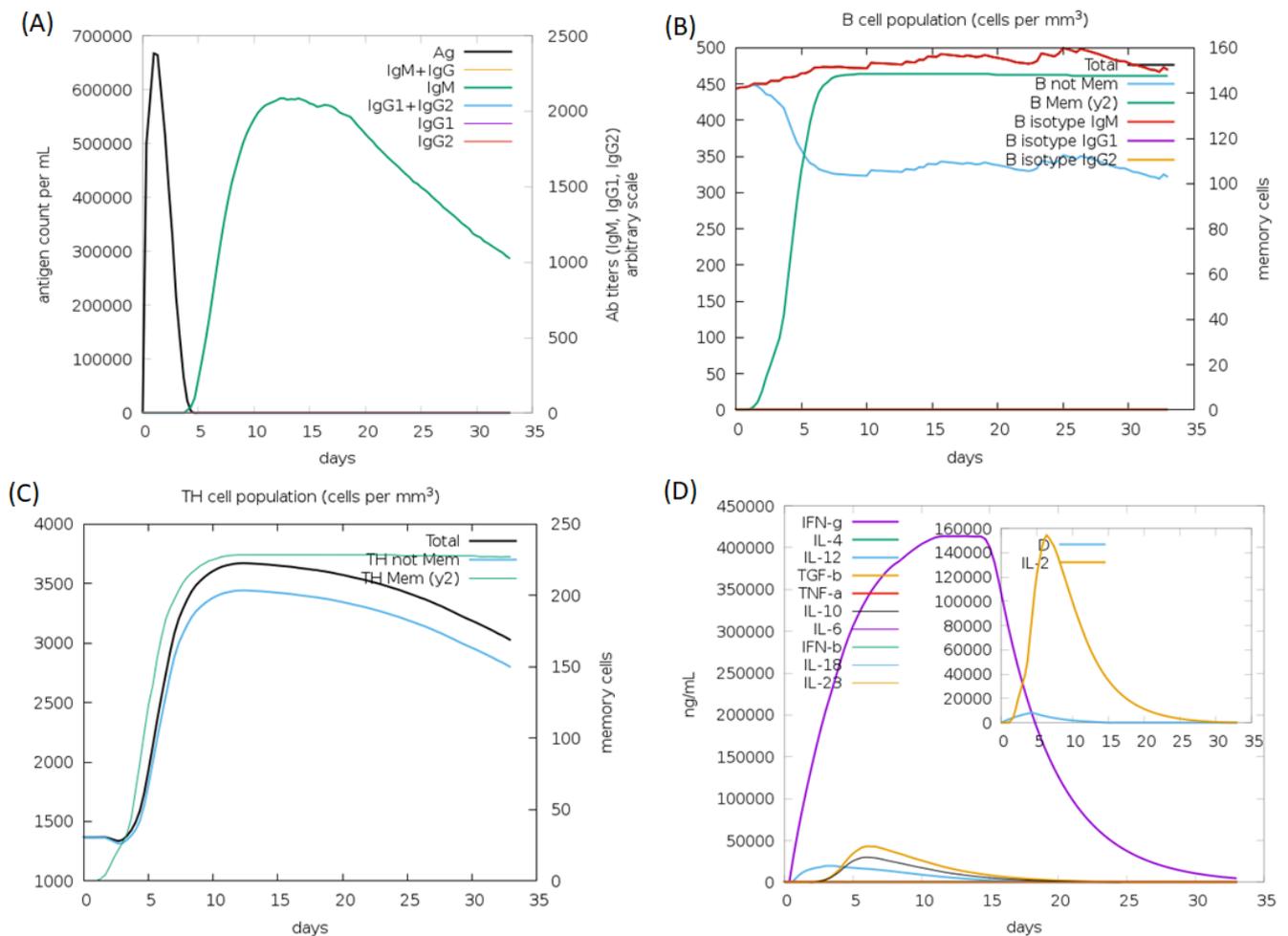


**Figure 3.** (A) Cartoon representation and (B) surface view of the structure and interactions of the docking complex (TLR3-designed vaccine). TLR3 is shown in green whereas designed vaccine is shown in red. (C) The dotted lines in white represent the polar interactions between the TLR3 and designed vaccine. (D) Ramachandran plot of the docking complex as revealed by PROCHECK server analysis.

### 3.10. *In Silico* Immunogenicity of the Designed Vaccine

The *in silico* host immune responses to our designed multi-epitope vaccine are shown in Figure 4. The immune responses after *in silico* stimulation included high IgM levels and

increased lymphocyte populations. Additionally, robust cytokine and interleukin responses were observed.



**Figure 4.** C-ImmSim server prediction results of the immune response against the designed multi-epitope vaccine; (A) Antigen and immunoglobulins; (B) B lymphocyte population; (C) Helper T cell population; (D) Induced levels of interleukins and cytokines.

#### 4. Discussion

SARS-CoV-2 has infected millions of people worldwide and is now considered a public health emergency. The development of efficient vaccines and drugs against COVID-19 disease is urgently needed to get rid of this pandemic. Several efforts are being carried out globally for the development of efficient and safe vaccines for this disease. Bioinformatics-based approaches offer a novel *in silico* methodology for the identification of potential T and B cell epitopes in the viral protein sequences. These epitopes are useful for the development of effective multi-epitope-based vaccines. Currently, several vaccine design efforts have been reported in the literature to address the challenge of vaccine construction for SARS-CoV-2 [13–25]. In most of the available studies, the immunoinformatics-based protocols were applied for the prediction of important B and T cell epitopes in various structural and non-structural proteins of SARS-CoV-2. The predicted epitopes were screened, and the selected ones were joined together to develop a multi-epitope vaccine. Here, we employed the immunoinformatic approach for the prediction of potential B and T cell epitopes in the RdRp protein of SARS-CoV-2. After scanning the primary amino acid sequence of RdRp protein using different high efficiency immunoinformatics tools, various B and T cell epitopes were identified. After proper screening of these epitopes using suitable bioin-

formatics servers, we identified and selected 3 B cell epitopes (Table 1), 10 HTL epitopes (Table 2), and 18 CTL epitopes (Table 3), which are highly antigenic and do not possess any toxicity or allergenicity (see Supplementary Materials). All the selected HTL epitopes are potential IFN- $\gamma$  inducers as revealed by the IFNepitope server. These selected peptide epitopes need further investigation under in vitro and in vivo conditions to confirm our in silico observations. Moreover, as reported in the literature, different peptide epitopes capable of activating several HLA restricted specific T cells can be combined with B cell epitopes to form a common vaccine formulation. Epitope-based vaccines are safe and have been reported to induce more potent immune responses than the parallel whole protein vaccines [55]. Accordingly, we joined our screened and selected B and T cell epitomes through different linkers to make the multi-epitope vaccine construct (Supplementary Figure S1). The multi-epitope vaccines are easy to produce in bulk using the codon optimization and gene cloning techniques. Since peptide vaccines are known to produce low immune responses in the absence of an efficient adjuvant, we linked the TLR-3 agonist  $\beta$ -defensin as an adjuvant to the multi-epitope vaccine construct. We reported here some additional in silico analysis, including the characterization and modeling of the vaccine construct and the validation of the generated vaccine model. Further, the molecular docking of the designed vaccine model with the TLR-3 structural model and the desired interaction between the two was also analyzed and reported (Figure 3). Moreover, the in silico host immune responses to our designed multi-epitope vaccine were observed to be quite significant (Figure 4). Our results suggest that this vaccine construct containing the immunogenic B and T cell epitopes of RdRp can be considered for the ongoing development of peptide-based vaccines against COVID-19. However, further in vitro and in vivo investigations are required to ensure the real potential of our designed multi-epitope vaccine to combat COVID-19.

## 5. Conclusions

In this study, the B and T cell epitopes in RNA-dependent RNA polymerase of SARS-CoV-2 were identified and screened using different web-based bioinformatics servers. The selected non-allergic and non-toxic epitopes with high antigenicity scores were joined together through linkers to make the multi-epitope vaccine construct. The vaccine construct was modeled, and the 3D model was docked with TLR3 using efficient docking servers. Stable molecular interaction between the vaccine model and TLR3 was observed in the docking complex. Further, the docking complex was validated by the PROCHEK server, and the favorable Ramachandran plot of the complex confirmed its stability and flexibility.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/vaccines10101660/s1>, Table S1. B-cell epitopes predicted using the BCPREDS server along with their VaxiJen score. Table S2. Top 40 HTL epitopes of RdRp predicted using IEDB server with their VaxiJen score. Table S3. CTL epitopes predicted using NetCTL 1.2 server with their VaxiJen score and ToxinPred server report. Table S4. ToxinPred report for CTL epitopes having VaxiJen score above the threshold of 0.4. Figure S1. The designed vaccine construct sequence. The sequence of epitopes in the designed vaccine is: Adjuvant-CTL epitopes-HTL epitopes-B cell epitopes. The first 68 amino acids are occupied by the adjuvant,  $\beta$ -defensin 1 (Blue color), next 5 amino acids are of EAAAK linker (Orange color) and the amino acids from 74–286 constitute the CTL epitopes (Red color) whereas the amino acids from 287–486 and from 487–540 correspond to HTL (Green color) and B-cell epitopes (Black color), respectively. All the linkers, except EAAAK are represented in purple color. Figure S2. Molecular dynamics simulations of the docked complex (TLR3-Designed vaccine). (A) The root-mean-square fluctuation (RMSF) values of the amino acid residues of Chain A (TLR3) of the docked complex. (B) RMSF values of the amino acid residues of Chain B (designed vaccine) of the docked complex.

**Author Contributions:** Conceptualization, S.A.M. and B.A.; methodology, S.A.M., A.A.B.D. and S.B.; software, S.A.M.; validation, S.A.M., Y.M. and A.A.O.; formal analysis, S.A.M.; investigation, M.A., B.A., A.A.B.D., S.B., Y.M. and A.A.O.; resources, S.A.M., M.A., S.B., Y.M. and S.A.A.; data curation, S.A.M. and Y.M.; writing—original draft preparation, S.A.M.; writing—review and editing, S.A.M., M.A., B.A., S.B., Y.M., S.A.A. and A.A.O.; visualization, S.A.M.; supervision, M.A., B.A., A.A.B.D. and

S.A.A.; project administration, S.A.M., M.A., A.A.B.D., S.B. and S.A.A.; funding acquisition, M.A., B.A., A.A.B.D. and S.B.; All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Deanship of Scientific Research at Majmaah University (Project number: R-2022-293).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** This study was financially supported by the Deanship of Scientific Research (Project number: R-2022-293) at Majmaah University, Al Majma'ah, Saudi Arabia.

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

## Abbreviations

CTL: Cytotoxic T-lymphocytes; HTL: Helper T-lymphocytes; IEDB: Immune Epitope Database; IFN- $\gamma$ : Interferon gamma; MHC: Major histocompatibility complex; MERS: Middle East Respiratory Syndrome; SARS: Severe Acute Respiratory Syndrome; TLR3: Toll-like receptor 3; WHO: World Health Organization.

## References

1. Tamalika, K.; Utkarsh, N.; Srijita, B.; Debashrito, D.; Filippo, C.; Mueller, D.M.; Srivastava, A.P. A candidate multi-epitope vaccine against SARS-CoV-2. *Sci. Rep.* **2020**, *10*, 10895.
2. Hui, D.S.; Azhar, E.I.; Madani, T.A.; Ntoumi, F.; Kock, R.; Dar, O.; Ippolito, G.; Mchugh, T.D.; Memish, Z.A.; Drosten, C.; et al. The continuing 2019-nCoV epidemic threat of novel coronaviruses to global health—The latest 2019 novel coronavirus outbreak in Wuhan, China. *Int. J. Infect. Dis.* **2020**, *91*, 264–266. [[CrossRef](#)] [[PubMed](#)]
3. Choudhary, S.; Malik, Y.S.; Tomar, S. Identification of SARS-CoV-2 Cell Entry Inhibitors by Drug Repurposing Using in silico Structure-Based Virtual Screening Approach. *Front. Immunol.* **2020**, *11*, 1664. [[CrossRef](#)] [[PubMed](#)]
4. Bergmann, C.C.; Lane, T.E.; Stohlman, S.A. Coronavirus infection of the central nervous system: Host–virus stand-off. *Nat. Rev. Microbiol.* **2006**, *4*, 121–132. [[CrossRef](#)] [[PubMed](#)]
5. Yi, Y.; Lagniton, P.N.; Ye, S.; Li, E.; Xu, R.H. COVID-19: What has been learned and to be learned about the novel coronavirus disease. *Int. J. Biol. Sci.* **2020**, *16*, 1753–1766. [[CrossRef](#)] [[PubMed](#)]
6. Perlman, S. Netland J Coronaviruses post-SARS: Update on replication and pathogenesis. *Nat. Rev. Microbiol.* **2009**, *7*, 439–450. [[CrossRef](#)] [[PubMed](#)]
7. Luk, H.K.H.; Li, X.; Fung, J.; Lau, S.K.P.; Woo, P.C.Y. Molecular epidemiology, evolution and phylogeny of SARS coronavirus. *Infect. Genet. Evol.* **2019**, *71*, 21–30. [[CrossRef](#)] [[PubMed](#)]
8. Coronavirinae in ViralZone. Available online: <https://viralzone.expasy.org/785> (accessed on 9 November 2020).
9. Kirchdoerfer, R.N.; Ward, A.B. Structure of the SARS-CoV nsp12 polymerase bound to nsp7 and nsp8 co-factors. *Nat. Commun.* **2019**, *10*, p2342. [[CrossRef](#)] [[PubMed](#)]
10. Ahn, D.G.; Choi, J.K.; Taylor, D.R.; Oh, J.W. Biochemical characterization of a recombinant SARS coronavirus nsp12 RNA-dependent RNA polymerase capable of copying viral RNA templates. *Arch. Virol.* **2012**, *157*, 2095–2104. [[CrossRef](#)]
11. Subissi, L.; Posthuma, C.C.; Collet, A.; Zevenhoven-Dobbe, J.C.; Gorbalenya, A.E.; Decroly, E.; Snijder, E.J.; Canard, B.; Imbert, I. One severe acute respiratory syndrome coronavirus protein complex integrates processive RNA polymerase and exonuclease activities. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, E3900–E3909. [[CrossRef](#)] [[PubMed](#)]
12. Chauhan, V.; Rungta, T.; Goyal, K.; Singh, M.P. Designing a multi-epitope based vaccine to combat Kaposi Sarcoma utilizing immunoinformatics approach. *Sci. Rep.* **2019**, *9*, 2517. [[CrossRef](#)] [[PubMed](#)]
13. Lim, H.X.; Lim, J.; Jazayeri, S.D.; Poppema, S.; Poh, C.L. Development of multi-epitope peptide-based vaccines against SARS-CoV-2. *Biomed. J.* **2021**, *44*, 18–30. [[CrossRef](#)] [[PubMed](#)]
14. Kiyotani, K.; Toyoshima, Y.; Nemoto, K.; Nakamura, Y. Bioinformatic prediction of potential T cell epitopes for SARS-Cov-2. *J. Hum. Genet.* **2020**, *65*, 569–575. [[CrossRef](#)] [[PubMed](#)]
15. Baruah, V.; Bose, S. Immunoinformatics-aided identification of T cell and B cell epitopes in the surface glycoprotein of 2019- nCoV. *J. Med. Virol.* **2020**, *92*, 495–500. [[CrossRef](#)] [[PubMed](#)]
16. Sahoo, B.; Kant, K.; Rai, N.K.; Chaudhary, D.K. Identification of T-cell epitopes in proteins of novel human coronavirus, SARS-Cov-2 for vaccine development. *Int. J. Appl. Biol. Pharm.* **2020**, *11*, 37–45.
17. Parvez, S.; Preeti, S. Prediction of T and B cell epitopes in the proteome of SARS-CoV-2 for potential use in diagnostics and vaccine design. *ChemRxiv* **2020**, preprint. [[CrossRef](#)]

18. Seema, M. T Cell Epitope-Based Vaccine Design for Pandemic Novel Coronavirus 2019-nCoV. *ChemRxiv* **2020**, preprint. [[CrossRef](#)]
19. Grifoni, A.; Sidney, J.; Zhang, Y.; Scheuermann, R.H.; Peters, B.; Sette, A. A sequence homology and bioinformatic approach can predict candidate targets for immune responses to SARS-CoV-2. *Cell Host Microbe* **2020**, *27*, 671–680.e2. [[CrossRef](#)]
20. Fast, E.; Chen, B. Potential T-cell and B-cell epitopes of 2019-nCoV. *BioRxiv* **2020**. [[CrossRef](#)]
21. Smith, T.R.F.; Patel, A.; Ramos, S.; Elwood, D.; Zhu, X.; Yan, J.; Gray, E.N.; Walker, S.N.; Schultheis, K.; Purwar, M.; et al. Immunogenicity of a DNA vaccine candidate for COVID-19. *Nat. Commun.* **2020**, *11*, 2601. [[CrossRef](#)]
22. Kalita, P.; Padhi, A.K.; Zhang, K.Y.J.; Tripathi, T. Design of a peptide-based subunit vaccine against novel coronavirus SARS-CoV-2. *Microb. Pathog.* **2020**, *145*, 104236. [[CrossRef](#)] [[PubMed](#)]
23. Bhattacharya, M.; Sharma, A.R.; Patra, P.; Ghosh, P.; Sharma, G.; Patra, B.C.; Lee, S.; Chakraborty, C. Development of epitope-based peptide vaccine against novel coronavirus 2019 (SARS-COV-2): Immunoinformatics approach. *J. Med. Virol.* **2020**, *92*, 618–631. [[CrossRef](#)] [[PubMed](#)]
24. Saha, R.; Prasad, B.V.L.S. *In silico* approach for designing of a multi-epitope based vaccine against novel Coronavirus (SARS-COV-2). *BioRxiv* **2020**, *9*, 10–16.
25. Ojha, R.; Gupta, N.; Naik, B.; Singh, S.; Verma, V.K.; Prusty, D.; Prajapati, V.K. High throughput and comprehensive approach to develop multiepitope vaccine against minacious COVID-19. *Eur. J. Pharm. Sci.* **2020**, *151*, 105375. [[CrossRef](#)] [[PubMed](#)]
26. Doytchinova, I.A.; Flower, D.R. VaxiJen: A server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinform.* **2007**, *8*, 4. [[CrossRef](#)] [[PubMed](#)]
27. Nielsen, M.; Lund, O. NN-align An artificial neural network-based alignment algorithm for MHC class II peptide binding prediction. *BMC Bioinform.* **2009**, *10*, 296. [[CrossRef](#)] [[PubMed](#)]
28. Larsen, M.V.; Lundegaard, C.; Lamberth, K.; Buus, S.; Lund, O.; Nielsen, M. Large-scale validation of methods for cytotoxic T-lymphocyte epitope prediction. *BMC Bioinform.* **2007**, *8*, 424. [[CrossRef](#)] [[PubMed](#)]
29. Dhanda, S.K.; Vir, P.; Raghava, G.P. Designing of interferon-gamma inducing MHC class-II binders. *Biol. Direct.* **2013**, *8*, 30. [[CrossRef](#)] [[PubMed](#)]
30. Sette, A.; Sidney, J. Nine major HLA class I supertypes account for the vast preponderance of HLA-A and -B polymorphism. *Immunogenetics* **1999**, *50*, 201–212. [[CrossRef](#)] [[PubMed](#)]
31. Shen, W.J.; Zhang, X.; Zhang, S.; Liu, C.; Cui, W. The utility of supertype clustering in prediction for class II MHC-peptide binding. *Molecules* **2018**, *23*, 3034. [[CrossRef](#)] [[PubMed](#)]
32. Nosrati, M.; Hajizade, A.; Nazarian, S.; Amani, J.; Vansofla, A.N.; Tarverdizadeh, Y. Designing a multi-epitope vaccine for cross-protection against Shigella spp: An immunoinformatics and structural vaccinology study. *Mol. Immunol.* **2019**, *116*, 106–116. [[CrossRef](#)] [[PubMed](#)]
33. Dimitrov, I.; Flower, D.R.; Doytchinova, I. AllerTOP—A server for in silico prediction of allergens. *BMC Bioinform.* **2013**, *14*, S4. [[CrossRef](#)] [[PubMed](#)]
34. Kumar, A.; Kumar, P.; Saumya, K.U.; Kapuganti, S.K.; Bhardwaj, T.; Giri, R. Exploring the SARS-CoV-2 structural proteins for multi-epitope vaccine development: An in-silico approach. *Expert Rev. Vaccines* **2020**, *19*, 887–898. [[CrossRef](#)] [[PubMed](#)]
35. Magnan, C.N.; Zeller, M.; Kayala, M.A.; Vigil, A.; Randall, A.; Felgner, P.L.; Baldi, P. High-throughput prediction of protein antigenicity using protein microarray data. *Bioinformatics* **2010**, *26*, 2936–2943. [[CrossRef](#)]
36. Dimitrov, I.; Naneva, L.; Doytchinova, I.; Bangav, I. AllergenFP: Allergenicity prediction by descriptor fingerprints. *Bioinformatics* **2013**, *30*, 846–851. [[CrossRef](#)]
37. Gasteiger, E.; Hoogland, C.; Gattiker, A.; Duvaud, S.; Wilkins, M.R.; Appel, R.D.; Bairoch, A. Protein Identification and Analysis Tools on the ExPASy Server. In *The Proteomics Protocols Handbook*; Springer Protocols Handbooks; Walker, J.M., Ed.; Humana Press: Totowa, NJ, USA, 2005; pp. 571–607. [[CrossRef](#)]
38. Jones, D.T. Protein secondary structure prediction based on position-specific scoring matrices. *J. Mol. Biol.* **1999**, *292*, 195–202. [[CrossRef](#)] [[PubMed](#)]
39. Buchan, D.W.A.; Jones, D.T. The PSIPRED protein analysis workbench: 20 years on. *Nucleic Acids Res.* **2019**, *47*, W402–W407. [[CrossRef](#)] [[PubMed](#)]
40. Obradovic, Z.; Peng, K.; Vucetic, S.; Radivojac, P.; Dunker, A.K. Exploiting heterogeneous sequence properties improves prediction of protein disorder. *Proteins* **2005**, *61* (Suppl. 7), 176–182. [[CrossRef](#)]
41. Obradovic, Z.; Peng, K.; Vucetic, S.; Radivojac, P.; Brown, C.J.; Dunker, A.K. Predicting intrinsic disorder from amino acid sequence. *Proteins* **2003**, *53* (Suppl. 6), 566–572. [[CrossRef](#)]
42. Xue, B.; Dunbrack, R.L.; Williams, R.W.; Dunker, A.K.; Uversky, V.N. PONDR-FIT: A meta-predictor of intrinsically disordered amino acids. *Biochim. Biophys. Acta* **2010**, *1804*, 996–1010. [[CrossRef](#)]
43. Mészáros, B.; Erdos, G.; Dosztányi, Z. IUPred2A: Context-dependent prediction of protein disorder as a function of redox state and protein binding. *Nucleic Acids Res.* **2018**, *46*, W329–W337. [[CrossRef](#)]
44. Gadhav, K.; Gehi, B.R.; Kumar, P.; Xue, B.; Uversky, V.N.; Giri, R. The dark side of Alzheimer’s disease: Unstructured biology of proteins from the amyloid cascade signaling pathway. *Cell Mol. Life Sci.* **2020**, *77*, 4163–4208. [[CrossRef](#)]
45. Kumar, A.; Kumar, P.; Kumari, S.; Uversky, V.N.; Giri, R. Folding and structural polymorphism of p53 C-terminal domain: One peptide with many conformations. *Arch. Biochem. Biophys.* **2020**, *684*, 108342. [[CrossRef](#)]
46. Kumar, D.; Singh, A.; Kumar, P.; Uversky, V.N.; Rao, C.D.; Giri, R. Understanding the penetrance of intrinsic protein disorder in rotavirus proteome. *Int. J. Biol. Macromol.* **2020**, *144*, 892–908. [[CrossRef](#)]

47. Yang, J.; Yan, R.; Roy, A.; Xu, D.; Poisson, J.; Zhang, Y. The I-TASSER Suite: Protein structure and function prediction. *Nat. Methods* **2015**, *12*, 7–8. [[CrossRef](#)]
48. Laskowski, R.A.; MacArthur, M.W.; Moss, D.S.; Thornton, J.M. PROCHECK: A program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **1993**, *26*, 283–291. [[CrossRef](#)]
49. Tovchigrechko, A.; Vakser, I.A. GRAMM-X public web server for protein-protein docking. *Nucleic Acids Res.* **2006**, *34*, W310–W314. [[CrossRef](#)]
50. Tovchigrechko, A.; Vakser, I.A. Development and testing of an automated approach to protein docking. *Proteins* **2005**, *60*, 296–301. [[CrossRef](#)]
51. DeLano, W.L. Pymol: An open-source molecular graphics tool. *CCP4 Newsl. Protein Crystallogr.* **2002**, *40*, 82–92.
52. Tahir ul Qamar, M.; Shahid, F.; Aslam, S.; Ashfaq, U.A.; Aslam, S.; Fatima, I.; Fareed, M.M.; Zohaib, A.; Chen, L.-L. Reverse vaccinology assisted designing of multiepitope-based subunit vaccine against SARS-CoV-2. *Infect. Dis. Poverty* **2020**, *9*, 132. [[CrossRef](#)]
53. Krieger, E.; Joo, K.; Lee, J.; Lee, J.; Raman, S.; Thompson, J.; Tyka, M.; Baker, D.; Karplus, K. Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: Four approaches that performed well in CASP8. *Proteins* **2009**, *77*, 114–122. [[CrossRef](#)] [[PubMed](#)]
54. Messaoudi, A.; Belguith, H.; Ben Hamida, J. Homology modeling and virtual screening approaches to identify potent inhibitors of VEB-1 beta-lactamase. *Theor. Biol. Med. Model* **2013**, *10*, 22. [[CrossRef](#)] [[PubMed](#)]
55. Kao, D.J.; Hodges, R.S. Advantages of a synthetic peptide immunogen over a protein immunogen in the development of an anti-pilus vaccine for *Pseudomonas aeruginosa*. *Chem. Biol. Drug Des.* **2009**, *74*, 33–42. [[CrossRef](#)] [[PubMed](#)]