



THE AGA KHAN UNIVERSITY

eCommons@AKU

Department of Biological & Biomedical
Sciences

Medical College, Pakistan

6-26-2021

Multi-subunit SARS-CoV-2 vaccine design using evolutionarily conserved T- and B- cell epitopes

Burkitkan Akbay

Syed Hani Abidi

Mahmoud A A. Ibrahim

Zhussipbek Mukhatayev

Syed Ali

Follow this and additional works at: https://ecommons.aku.edu/pakistan_fhs_mc_bbs



Part of the [Influenza Virus Vaccines Commons](#), [Life Sciences Commons](#), and the [Virus Diseases Commons](#)

Article

Multi-Subunit SARS-CoV-2 Vaccine Design Using Evolutionarily Conserved T- and B- Cell Epitopes

Burkitkan Akbay ^{1,†} , Syed Hani Abidi ^{2,*,†} , Mahmoud A. A. Ibrahim ³ , Zhussipbek Mukhatayev ¹ and Syed Ali ^{1,*}

¹ Department of Biomedical Sciences, Nazarbayev School of Medicine, Nazarbayev University, Nur-Sultan 010000, Kazakhstan; burkitkan.akbay@nu.edu.kz (B.A.); zhussipbek.mukhatayev@nu.edu.kz (Z.M.)

² Department of Biological and Biomedical Sciences, Aga Khan University, Karachi 74800, Pakistan

³ Computational Chemistry Laboratory, Chemistry Department, Faculty of Science, Minia University, Minia 61519, Egypt; m.ibrahim@compchem.net

* Correspondence: m.haniabidi@gmail.com (S.H.A.); syed.ali@nu.edu.kz (S.A.); Tel.: +7-7054231132 (S.A.)

† These authors contributed equally to this work.

Abstract: The SARS-CoV-2 pandemic has created a public health crisis worldwide. Although vaccines against the virus are efficiently being rolled out, they are proving to be ineffective against certain emerging SARS-CoV-2 variants. The high degree of sequence similarity between SARS-CoV-2 and other human coronaviruses (HCoV) presents the opportunity for designing vaccines that may offer protection against SARS-CoV-2 and its emerging variants, with cross-protection against other HCoVs. In this study, we performed bioinformatics analyses to identify T and B cell epitopes originating from spike, membrane, nucleocapsid, and envelope protein sequences found to be evolutionarily conserved among seven major HCoVs. Evolutionary conservation of these epitopes indicates that they may have critical roles in viral fitness and are, therefore, unlikely to mutate during viral replication thus making such epitopes attractive candidates for a vaccine. Our designed vaccine construct comprises of twelve T and six B cell epitopes that are conserved among HCoVs. The vaccine is predicted to be soluble in water, stable, have a relatively long half-life, and exhibit low allergenicity and toxicity. Our docking results showed that the vaccine forms stable complex with toll-like receptor 4, while the immune simulations predicted that the vaccine may elicit strong IgG, IgM, and cytotoxic T cell responses. Therefore, from multiple perspectives, our multi-subunit vaccine design shows the potential to elicit a strong immune-protective response against SARS-CoV-2 and its emerging variants while carrying minimal risk for causing adverse effects.

Keywords: human coronaviruses; MERS; SARS-CoV-2; SARS-CoV; epitope; vaccine



Citation: Akbay, B.; Abidi, S.H.; Ibrahim, M.A.A.; Mukhatayev, Z.; Ali, S. Multi-Subunit SARS-CoV-2 Vaccine Design Using Evolutionarily Conserved T- and B- Cell Epitopes. *Vaccines* **2021**, *9*, 702. <https://doi.org/10.3390/vaccines9070702>

Academic Editors: Stephen A. Morris and Camila Coelho

Received: 6 May 2021

Accepted: 14 June 2021

Published: 26 June 2021

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



Copyright: © 2021 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

The novel severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has caused a worldwide pandemic that continues to negatively impact the global economy, social dynamics, and health care systems [1]. SARS-CoV-2 has been identified as the seventh coronavirus that can infect humans [2]. Although coronaviruses have infected the human population repeatedly over the past century [3], only a few have caused severe large-scale outbreaks. Prior to SARS-CoV-2, the SARS-CoV outbreak occurred in southern China in November 2002, followed by the Middle East respiratory syndrome (MERS-CoV) outbreak in Saudi Arabia and South Korea, respectively, in 2012 and 2015. SARS-CoV and MERS-CoV have an estimated mortality rate of ~9.6% and 35%, respectively [4,5]. Moreover, four other human coronaviruses (HCoVs) associated with mild upper respiratory symptoms, have previously been identified, i.e., HCoV-NL63, HCoV-229E, HCoV-OC43, and HCoV-HKU1 [6–8].

The genomes of SARS-CoV and MERS-CoV are, respectively, 79% and 50% similar to SARS-CoV-2 [9]. Out of these HCoVs, SARS-CoV and SARS-CoV-2 not only share a phylogenetic relationship, but also exhibit a high level of similarity in the cell entry mechanism and the use of human cell receptors [10,11].

The urgency for having an effective vaccine against SARS-CoV-2 has given rise to fervent efforts toward vaccine design worldwide [12,13]. Over 200 candidate vaccines have been developed, and several of them are currently authorized for use in different parts of the world and have shown varying efficacy [14]. Two mRNA-based vaccines, BNT162b2 (Pfizer–BioNTech, New York, NY, USA and Mainz, Rhineland-Palatinate, Germany), mRNA-1273 (Moderna, Cambridge, MA, USA), showed up to 95% efficacy, the adenovirus vector-based Ad26.CoV2.S (Johnson & Johnson, New Brunswick, NJ, USA) presented 72% efficacy in the USA but 57% in South Africa, while the chimpanzee adenovirus vector-based ChAdOx (AZD1222) (AstraZeneca, Cambridge, UK) presented 70% efficacy. Similarly, heterologous recombinant adenovirus vector-based Gam-COVID-Vac (Sputnik V, Moscow, Russia) demonstrated 91% efficacy in the preliminary phase 3 trials, while the protein-based vaccine NVX-CoV237 (Novavax, Gaithersburg, MD, USA) was 89% efficient in the U.K. but 49% in South Africa [14–18]. Two inactivated virus vaccines, CoronaVac (Sinovac, Beijing, China) and BBIBP-CorV (Sinopharm, China National Pharmaceutical Group, Beijing, China), passed through phase 3 clinical trials and were reported to show from 50% to 91% efficacy. The data, however, were not published at the time of writing of this manuscript. In all of the vaccines that have been designed so far, the antigenic target used was the spike protein, which is the most antigenic, highly variable, and prone to mutations [19,20]. Due to a high substitution/mutation rate and rapid transmissions, new variants of SARS-CoV-2 have been continuously emerging globally, posing threats to circumvent the efficacy of current vaccines, possibly leading to a novel-SARS-CoV-2-variant-associated surge in transmissions [15]. In fact, there have already been reports of the current vaccines losing effectiveness against the emerging SARS-CoV-2 variants [21,22]. Therefore, vaccines that can stay effective against the emerging strains while simultaneously exhibiting an acceptable safety profile are the need of the hour.

One possible way to design such vaccines is to include epitopes, aside from spike, from SARS-CoV-2 that are evolutionary conserved. Structural proteins of coronaviruses, including spike (S), envelope (E), membrane (M), and nucleocapsid (N), have been shown to elicit greater immunogenicity compared to those against the non-structural proteins [23]. A vaccine bearing a design that is based on conserved epitopes of structural proteins is likely to remain effective against emerging variants since mutations in conserved regions, as they incur a fitness cost to the virus, are not likely to occur. Corroborating that argument are several studies showing that individuals with previous exposure to other HCoVs exhibit high protection against SARS-CoV-2 [24], indicating that due to their functional significance, certain antigenic regions have remained conserved in all HCoVs.

In this study, using *in silico* analyses, we designed a multi-subunit chimeric SARS-CoV-2 vaccine comprising of CD8+ T and linear B cell epitopes originating from the regions of envelope (E), membrane (M), nucleocapsid (N), and spike (S) proteins found to be conserved among several human coronaviruses. Evolutionary conservation of these motifs implies their critical roles in viral fitness, indicating that they are unlikely to become mutated during viral replication, making them attractive candidates for a SARS-CoV-2 subunit vaccine that promises to provide coverage against multiple HCoVs, including SARS-CoV-2 and its emerging variants.

2. Materials and Methods

2.1. Nucleotide Sequence Retrieval and Phylogenetic Analysis

In the first step, we used E, N, M, and S reference sequences for all seven HCoVs, available at the NCBI Genbank database, namely, SARS-CoV-2 (NC_045512), SARS-CoV (NC_004718), MERS-CoV (NC_019843), HCoV-NL63 (NC_005831), HCoV-229E (NC_002645), HCoV-HKU1 (NC_006577), and HCoV-OC43 (NC_006213) to retrieve nucleotide sequences

for each HCoV exhibiting >90% sequence similarity to the reference sequences. The retrieved sequences were aligned using the online MAFFT alignment tool (<https://mafft.cbrc.jp/alignment/server/> (accessed on 23 May 2020)). In order to analyze the evolutionary relationship between SARS-CoV-2 with other HCoVs (SARS-CoV, MERS-CoV, HCoV-NL63, HCoV-229E, HCoV-OC43, and HKU-1), nucleotide alignment was used to construct a maximum-likelihood (ML) phylogenetic tree using IQ-TREE v1.6.12 [25]. The ML tree was constructed using the general time reversible (GTR) model of nucleotide substitution and gamma correction for among site rate variation. The ultrafast bootstrap method was used with 1000 replicates to infer support branching in the tree topology. The Newick file for the ML tree was visualized using FigTree v1.4.4.

2.2. Protein Sequence Retrieval and Vaccine Design Workflow

In the next step, E, N, M, and S reference sequences for all seven HCoVs (described above) were used in BLAST search (pBLAST), with filters of 90–100% percent identity and 100% query coverage, to retrieve E, N, M, and S protein sequences from HCoVs that are over 90% conserved. The total number of protein sequences from each HCoV used in our study are shown in Table 1.

Table 1. Number of protein sequences for envelope, membrane, nucleocapsid, and spike proteins from HCoVs. The table shows the number of sequences used in our study that gave 100% query coverage and >90% conservation between HCoVs.

HCoVs	Protein Sequences			
	Envelope	Membrane	Nucleocapsid	Spike
HCoV-NL63	2	11	36	51
HCoV-229E	5	6	51	71
HCoV-OC43	6	15	106	183
HCoV-HKU1	3	8	28	26
MERS-Cov	18	35	87	250
SARS-CoV	13	23	29	80
SARS-CoV-2	56	176	3082	1139

Subsequently, the E, N, M, and S protein sequences for each HCoV were separately aligned using an online MAFFT alignment tool [26], and each alignment was used to generate protein specific consensus sequences for each HCoV with the Consensus Maker tool (<https://www.hiv.lanl.gov/content/sequence/CONSENSUS/SimpCon.html> (accessed on 25 May 2020)). In the next step, protein-specific consensus sequences from the HCoVs were aligned again, and the conserved regions within the E, N, M, and S proteins from the seven HCoVs were identified using ESPript 3.x [27] and Jalview 2.11.1.3 [28] software. Additionally, to take into account the recently emerging mutations among SARS-CoV-2 variants, approximately 5000 whole genome sequences of three SARS-CoV-2 variants: B.1.1.7 (VOC-202012/01) identified in the U.K., 501Y.V2 (B.1.351) identified in the South Africa, and 501Y.V3 (P.1) identified in Brazil were retrieved from GSAID (<https://www.gisaid.org/> (accessed on 5 April 2021)), and aligned with MAFFT to generate consensus sequences for each of the variants using BioEdit [29]. Nucleotide consensus sequences were then translated into amino acid sequences using the ExPasy tool (<https://web.expasy.org/translate/> (accessed on 5 April 2021)). The resulting amino acid sequences were trimmed with reference to E, N, M, and S reference protein sequences in Aliview [30]. The protein-specific consensus sequences for each HCoV, including the SARS-CoV-2 variants, were used to predict B and T cell epitopes and to design a multi-subunit vaccine based on the selected epitopes as described below (Figure 1).

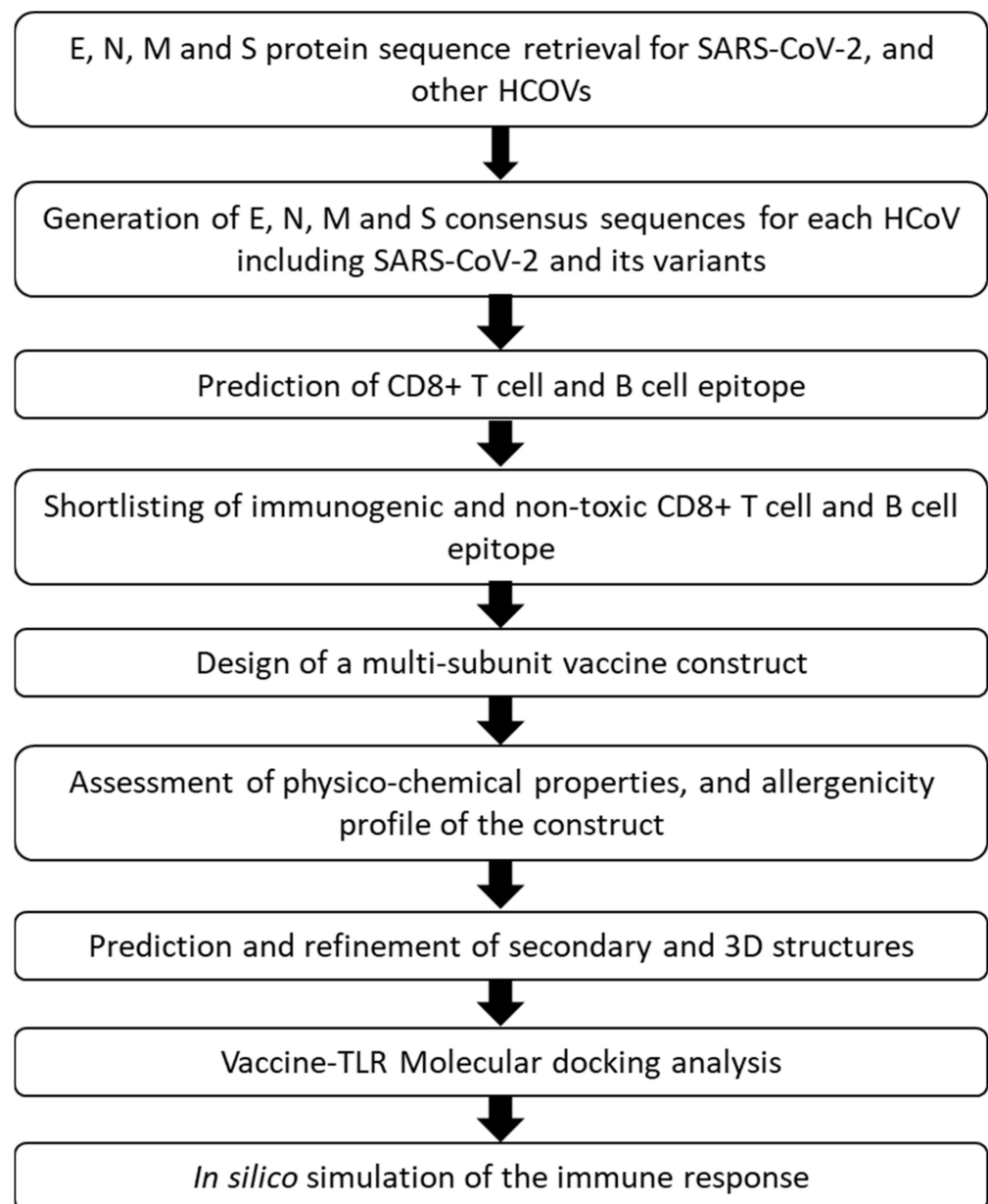


Figure 1. Vaccine design workflow.

2.3. *In Silico* CD8+ T Cell and B Cell Epitope Mapping, and Epitope Conservation Analysis

Using E, N, M, and S protein consensus sequences, HCoV-derived cytotoxic T cell epitopes were predicted using the web-based CTLPred tool [31] and a combined support vector machine and artificial neural network approach. Human leukocyte antigen (HLA) restriction was identified using the nHLAPred tool for each predicted T cell epitope [32]. The prediction of linear B cell epitopes was performed using the web-based ABCpred tool [33]. Epitopes located in the conserved regions between the 7 HCOVs or 3 closely related viruses, namely, SARS-CoV-2, SARS-CoV, and MERS-CoV, were selected. Predicted CD8+ T and B cell epitopes were compared to the experimentally validated SARS-CoV epitopes available in IEDB (<http://www.iedb.org/> (accessed on 10 April 2021)) database [34] and that had been reported by other studies.

2.4. Selection of Immunogenic and Non-Toxic Epitopes for Multi-Subunit Vaccine Design

The predicted epitopes were shortlisted after screening for immunogenicity and toxicity. VaxiJen v2.0 [35] was used to evaluate immunogenicity (a score >0.4 (default

value) was considered as antigenic), while ToxinPred [36] was used to evaluate the toxicity of the epitopes. Epitopes predicted to be immunogenic and non-toxic were selected to design the multi-subunit vaccine.

2.5. Design of a Multi-Subunit Vaccine

As indicated earlier, the conserved epitopes predicted to be immunogenic and non-toxic were used to design a multi-subunit vaccine (Supplementary Figure S1). For the vaccine design, the CD8+ T cell epitopes were linked using an AAY linker, and linear B cell epitopes were linked by an GPGPG linker [37]. At the C-terminal of the vaccine, a six-histidine tail was added while at the N-terminus, β -defensin amino acid sequence (UniProt ID: Q5U7J2) was added to the first epitope via an EAAAK linker. β -defensin acts as an adjuvant and mediates recruitment of naïve T cells and immature dendritic cells at the site of infection/immune activity [38].

2.6. Evaluation of Physico-Chemical Properties, and Allergenicity Profile of the Multi-Subunit Vaccine

Physico-chemical properties, such as half-life, instability index, aliphatic index, and grand average of hydrophobicity of the vaccine were determined using the ProtParam server (<http://web.expasy.org/protparam> (accessed on April 12 2021)), while the basic property of allergenicity of the multi-subunit vaccine was assessed using AllerTOP v. 2.0 server [39]. The immunogenicity/antigenicity of the multi-subunit vaccine was evaluated using the VaxiJen v2.0 server [35].

2.7. Prediction and Refinement of Secondary and 3D Structures

The secondary structure of the designed vaccine was predicted using the PSIPRED server [40], while the 3D structure was predicted using both the I-TASSER [41] and the Phyre2 server (using an intensive modelling approach) [42]. The best model was selected and refined/optimized using the Galaxyrefine server [43], while PROSA [44] and Ramachandran plot analyses were performed to analyze the quality of the refined structure. Solubility of the protein (vaccine) was determined using the Protein-Sol server [45].

2.8. Molecular Docking Analysis

Since Toll-like receptor 4 (TLR4) is capable of recognizing various exogenous and endogenous ligands such as viral protein, lipopolysaccharide, heat shock proteins, β -defensin, etc., which leads to dendritic cell maturation, and up-regulation of costimulatory molecules resulting in type 1 polarized adaptive immune response [46,47], we performed molecular docking between TLR4 and our vaccine-construct (containing β -defensin as an adjuvant). For molecular docking calculations, the crystal structure of the human TLR4 deposited in the Protein Data Bank with the PDB accession code of 4G8A [48] was retrieved and utilized as a template. For the preparation of the TLR4 crystal structure, all crystallographic water molecules and heteroatoms were removed. Moreover, mutated residues were reversed and refined. The protonation state of the TLR4 was then examined, and hydrogen atoms were consequently added [49]. Binding modes of the developed vaccine with the human TLR4 were then predicted using the PatchDock server [50]. All predicted docked complexes (exactly 1000 solutions) were then refined and rescored using the FireDock server [51,52]. The interactions between the vaccine and the TLR4 were resolved using the LigPlot + software [53].

2.9. In Silico Simulation of the Immune Response

The humoral and cellular immune response against the multi-subunit vaccine construct was simulated using the C-ImmSim tool [54]. For simulations, the following parameters were used: three injections at 4-week intervals; simulation steps = 1000; time to simulation = default; injection volume = vaccine only and no LPS. Additionally, we used

NetCTL [55] and NetChop [56,57] to predict proteasomal sites (using default a value of 0.5) and the possible proteolytic release of the epitopes from the vaccine construct.

3. Results

3.1. Phylogenetic and Sequence Similarity among the Human Coronaviruses

To evaluate the genetic similarity among the seven human coronaviruses, we constructed ML phylogenetic trees that revealed alpha-HCoVs, NL63 and 229E, clustered distantly from Beta-HCoVs, OC43 and HKU1. Importantly, SARS-CoV-2, SARS-CoV, and MERS-CoV clustered together, indicating a close phylogenetic relatedness between these viruses (Supplementary Figure S2).

In the next step, we identified conserved regions in the consensus sequences of the E, N, M, and S proteins from seven HCoVs (Figure 2A and Supplementary Figure S3) and three phylogenetically closely related HCoVs-SARS-CoV-2, SARS-CoV, and MERS-CoV (Figure 2B and Supplementary Figure S4) using ESPrnt 3.x and Jalview 2.11.1.3 software [28]. The results revealed several regions within N, M, and S proteins exhibiting more than 40% homology among the seven HCoVs, while no conserved regions were found for the E protein (Supplementary Figure S3), corroborating the findings of the phylogenetic analysis. Additionally, while the conservation was analyzed for SARS-CoV-2, SARS-CoV, and MERS-CoV, the results showed greater sequence conservation for E, N, M, and S compared to those in the seven hCoVs (Supplementary Figures S3 and S4).

In the next steps, focusing on these conserved regions, we predicted CD8+ T and B cell epitopes for each SARS-CoV-2 protein.

3.2. CD8+ T Cell Epitopes Originating in Conserved Regions of E, N, M, and S Proteins

CD8+ T cell epitopes were analyzed for E, N, M, and S proteins, focusing on the conserved regions of the seven HCoVs protein sequences.

No CD8+ T cell epitopes were identified in the conserved regions of E protein among all seven HCoVs (Supplementary Figure S3 and Table 2). However, when the analysis was carried out using E sequences from only SARS-CoV-2, SARS-CoV and MERS-CoV, two conserved CD8+ T cell epitopes were identified, each exhibiting ~44% and ~77% conservation (Table 3 and Supplementary Figure S4). These epitopes were found to be restricted by HLA-C*0401 HLA-A*0205, HLA-A*0301, HLA-A2, HLA-A3, and HLA-B*5301 (Table 3). Similarly, analysis of the M protein sequences showed the presence of one CD8+ T cell epitope with ~55% conservation among all seven HCoVs (Table 2), which was found to be restricted by HLA-A24, HLA-A*2402, and HLA-Cw*0401 (Table 2). Conservation of this epitope between SARS-CoV-2, SARS-CoV, and MERS-CoV was found to be 66% (Table 2 and Supplementary Figure S4).

Further analysis of SARS-CoV-2, SARS-CoV, and MERS-CoV epitopes revealed three conserved CD8+ T cell epitopes exhibiting ~44%, ~66%, and ~88% conservation between the three viruses (Table 3 and Supplementary Figure S4). These epitopes were found to be restricted by HLA-C*0401 and then by HLA -A*0203, -A*24, -B*2703, and -B*35 (Table 3).

Analysis of the N protein sequences showed one CD8+ T cell epitope with 44% conservation among all HCoVs (Table 2 and Supplementary Figure S3). This epitope was found to be restricted by HLA-Cw*0401, HLA-A*0202, HLA-A*0203, HLA-A*0205, HLA-A*0301, HLA-A*0401, HLA-B*35, and HLA-B*3501 (Table 2). Again, conservation of this epitope was increased to ~77% when the analysis was focused only on SARS-CoV-2, SARS-CoV, and MERS-CoV (Table 3 and Supplementary Figure S4). Seven additional conserved CD8+ T cell epitopes were identified for the N protein with ~77%, ~66%, and ~55% conservation (Table 3 and Supplementary Figure S4). Most of the conserved epitopes were found to be restricted by HLA-C*0401 and then by HLA-A*0201, HLA-A*2402, HLA-A*A2, HLA-A*6801, HLA-B*3501, HLA-B*51, and HLA-B*5301 (Table 3).

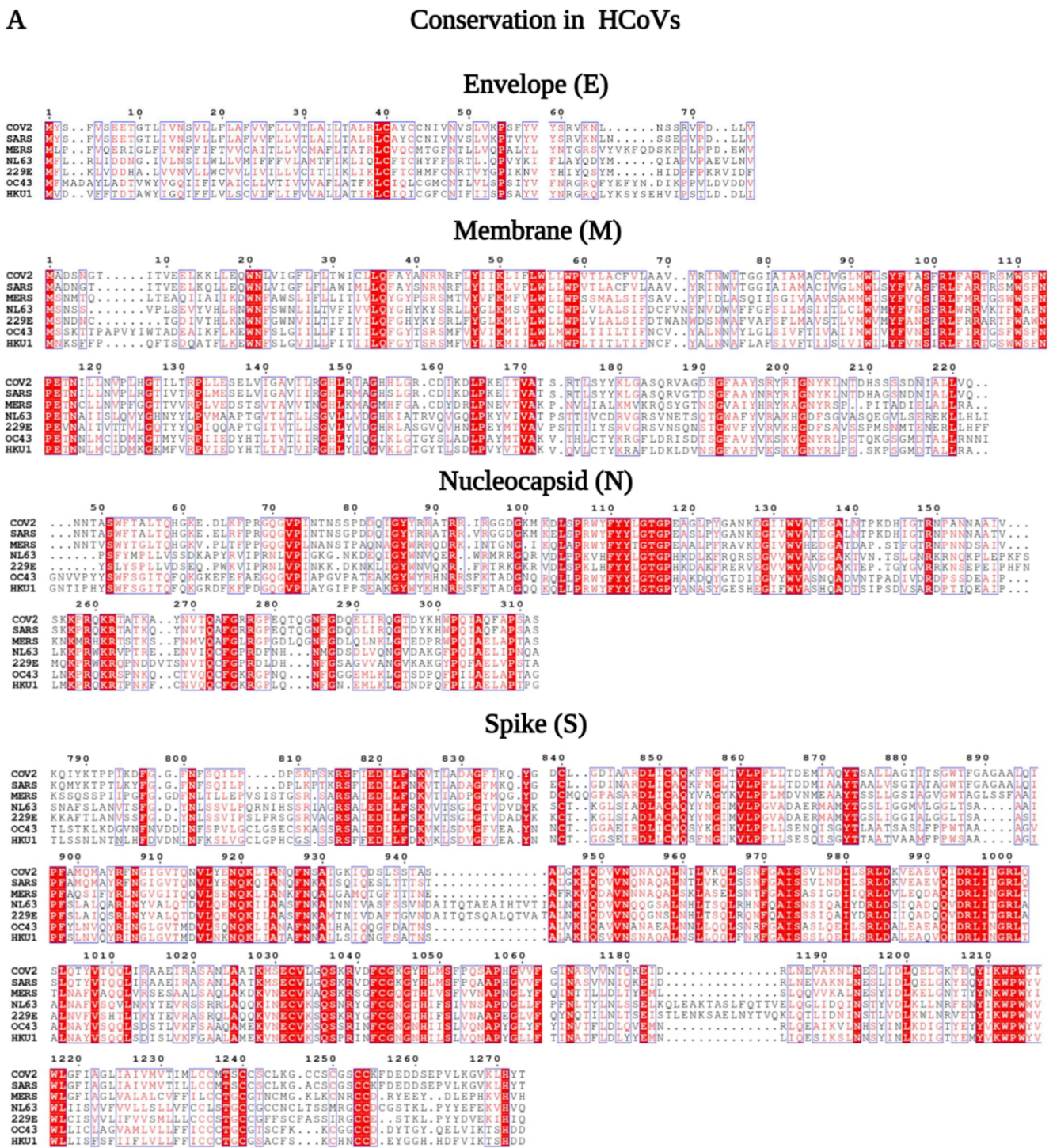


Figure 2. Cont.

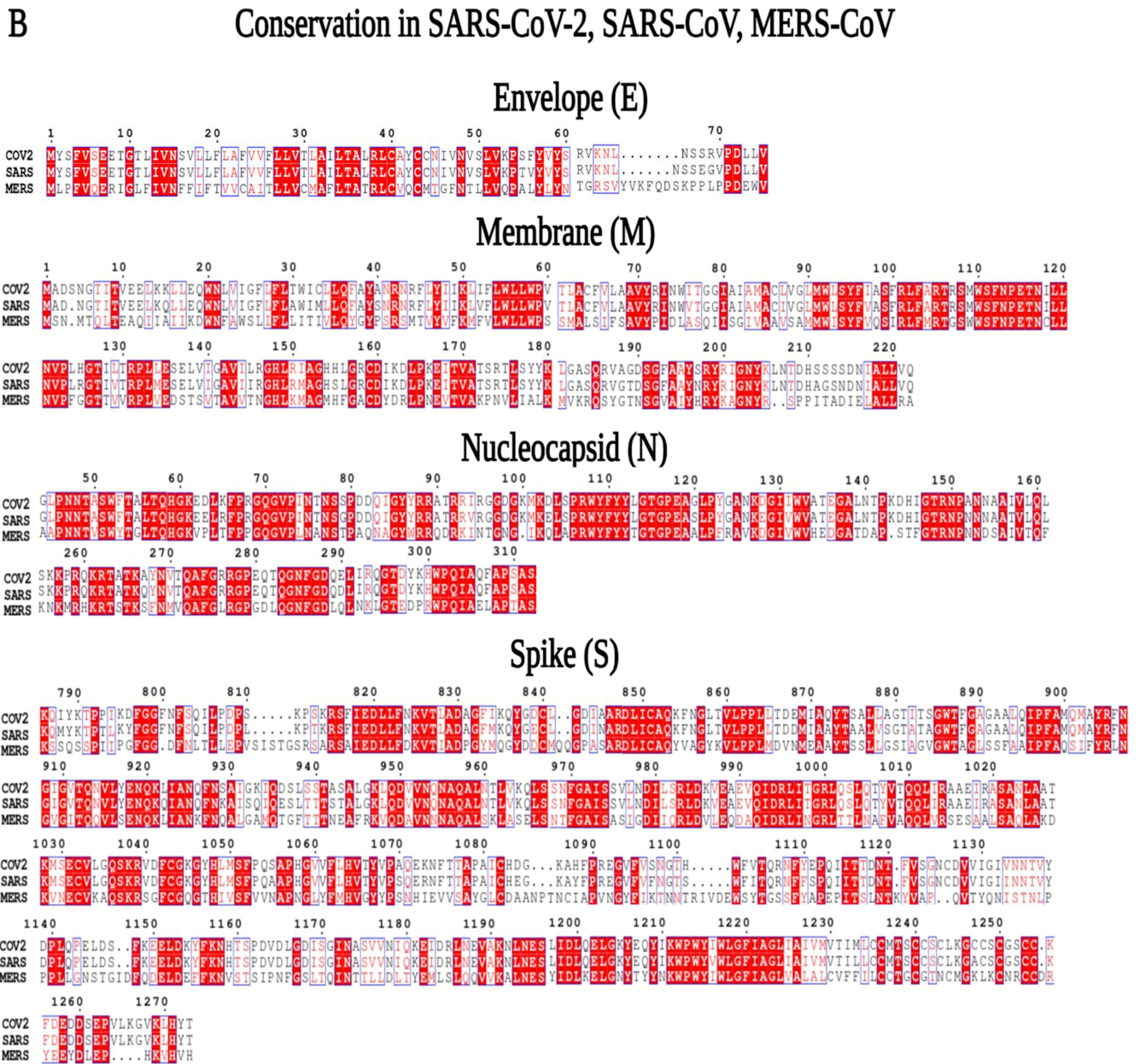


Figure 2. Conservation of envelope (E), membrane (M), nucleocapsid (N), and spike (S) protein sequences in seven HCoV and in SARS-CoV-2, SARS-CoV, and MERS-CoV. Conservation of E, M, N, and S proteins in (A) seven HCoV and (B) SARS-CoV-2 (COV2), SARS-CoV (SARS), and MERS-CoV (MERS). Only the regions showing conservancy are presented for N and S proteins (full length sequences are shown in Supplementary Figures). Images are shown from ESPrift results. Conserved regions are colored in red.

Table 2. Predicted CD8+ T cell epitopes found to be conserved in the seven HCoV. Conserved amino acids are shown in bold letters. HLA restriction for each epitope is also given.

Conserved CD8 + T Cell Epitopes in HCoVs															
Membrane															
Nº	SARS-COV-2	HLA-Restriction	SARS-COV	HLA-Restriction	MERS	HLA-Restriction	NL63	HLA-Restriction	229E	HLA-Restriction	OC43	HLA-Restriction	HKU1	HLA-Restriction	Conservation (%)
1	SYFIASFRL	HLA-A24,HLA-Cw*0401	SYFVASFRL	HLA-A24,HLA-A*2402,HLA-Cw*0401	SYFVQSIRL	HLA-Cw*0401	MYFVNSFRL	HLA-A24,HLA-Cw*0401	MYFANSFRL	HLA-A24,HLA-Cw*0401	VYFVNSIRL	HLA-A24,HLA-A*2402,HLA-Cw*0401	YFVNSIRLF	HLA-Cw*0401	55
Nucleocapsid															
1	SPRWYFYLL	HLA-Cw*0401	SPRWYFYLL	HLA-Cw*0401	APRWYFYTY	HLA-A*0202,HLA-A*0203,HLA-A*0205,HLA-A*0301,HLA-Cw*0401	PPKVHFYLL	HLA-A*0301,HLA-Cw*0401	SPKLHFYLL	HLA-B*3501,HLA-Cw*0401,HLA-B35	LPRWYFYLL	HLA-A*0301,HLA-Cw*0401,HLA-B35	LPRWYFYLL	HLA-A*0301,HLA-Cw*0401,HLA-B35	44
Spike															
1	YIKWPWYIW	HLA-A*0206,HLA-Cw*0401	YIKWPWYVW	HLA-A*0206,HLA-Cw*0401	YNKWPWYIW	HLA-A*0206,HLA-B*3501,HLA-B*5102,HLA-B*5103,HLA-B*5401,HLA-B7,HLA-Cw*0401,HLA-B35	YIKWPWVWV	HLA-A2,HLA-A*0206,HLA-Cw*0401	YIKWPWVWV	HLA-A2,HLA-A*0206,HLA-Cw*0401	YVKWPWYVW	HLA-A*0206,HLA-B*51,HLA-Cw*0401	YVKWPWYVW	HLA-A*0206,HLA-B*51,HLA-Cw*0401	66
2	QSAPHGVVF	HLA-Cw*0401,HLA-A*3301	QAAPHGVVF	HLA-Cw*0401,	VNAPNGLYF	HLA-B*5401,HLA-B*51,HLA-B*0702,HLA-Cw*0401	NSAPDGLLF	HLA-Cw*0401	NAAPEGLVF	HLA-Cw*0401	QNAPYGLYF	HLA-B*5401,HLA-Cw*0401,HLA-A*3301	QNAPYGLLF	HLA-B*5401,HLA-Cw*0401	44

Table 3. Predicted CD8+ T cell epitopes found to be conserved SARS-CoV-2, SARS-CoV, and MERS-CoV. Conserved amino acids are shown in bold letters. HLA restriction for each epitope is also given.

Conserved CD8 + T Cell Epitopes in SARS-CoV-2, SARS-CoV, and MERS-CoV							
Envelope							
Nº	SARS-CoV-2	HLA-resctriciton	SARS-CoV	HLA-resctriciton	MERS-CoV	HLA-resctriciton	Conservation (%)
1	AILTALRLC	HLA-Cw*0401	AILTALRLC	HLA-Cw*0401	AFLTATRLC	HLA-Cw*0401	77
2	SLVKPSFYV	HLA-Cw*0401	SLVKPTVYV	HLA-A*0205,HLA-A*0301,HLA-B*5301,HLA-Cw*0401	LLVQPALYL	HLA-A2,HLA-A3,HLA-A*0301,HLA-Cw*0401	44
Membrane							
1	SFNPETNIL	HLA-A*0203,HLA-Cw*0401,	SFNPETNIL	HLA-A*0203,HLA-Cw*0401,	SFNPETNCL	HLA-Cw*0401	88
2	SYFIASFRL	HLA-A24,HLA-Cw*0401	SYFVASFRL	HLA-A24,HLA-A*2402,HLA-Cw*0401	SYFVQSIRL	HLA-Cw*0401	66
3	KDLPKEITV	HLA-Cw*0401	KDLPKEITV	HLA-Cw*0401	DRLPNEVTV	HLA-B*51,HLA-Cw*0401	66
4	AGDSGFAAY	HLA-A1,HLA-Cw*0401,HLA-B35,HLA-B*2703	GTDSGFAAY	HLA-A1,HLA-A3,HLA-Cw*0401,HLA-B35,HLA-B*2703	GTNSGVAIY	HLA-A3,HLA-Cw*0401,HLA-B35,HLA-B44	44
Nucleocapsid							
1	SPRWYFYYL	HLA-Cw*0401	SPRWYFYYL	HLA-Cw*0401	APRWYFYYT	HLA-A*0202,HLA-A*0203,HLA-A*0205,HLA-A*0301,HLA-Cw*0401	77

Table 3. Cont.

Conserved CD8 + T Cell Epitopes in SARS-CoV-2, SARS-CoV, and MERS-CoV							
2	ASAFFGMSR	HLA-Cw*0401,HLA-A*6801	ASAFFGMSR	HLA-Cw*0401,HLA-A*6801	ASAFMGMSQ	HLA-A*1101,HLA-A3,HLA-A31,HLA-Cw*0401,HLA-A*3301,HLA-A*6801	77
3	ANKDGIWV	HLA-B*3501,HLA-B*5101,HLA-B*5301,HLA-B*5401,HLA-B*51,HLA-B*0702,HLA-Cw*0401,HLA-B35	ANKEGIVWV	HLA-B*3501,HLA-B*5101,HLA-B*5301,HLA-B*5401,HLA-B*51,HLA-B*0702,HLA-Cw*0401,HLA-B35	AVKDGIVWV	HLA-B*5301,HLA-B*51,HLA-B8,HLA-Cw*0401	66
4	QASSRSSSR	HLA-Cw*0401	QASSRSSSR	HLA-Cw*0401	QSSSRASSV	HLA-B*5301,HLA-Cw*0401	66
5	LLDRLNQL	HLA-A2,HLA-A*0201,HLA-A*0206,HLA-Cw*0401	LLDRLNQL	HLA-A2,HLA-A*0201,HLA-A*0206,HLA-Cw*0401	LYLDLLNRL	HLA-A2,HLA-A24,HLA-Cw*0401	66
6	VPINTNSSP	HLA-Cw*0401	VPINTNSGP		VPLNANSTP	HLA-Cw*0401	66
7	GYRRATRR	HLA-Cw*0401	GYRRATRR	HLA-Cw*0401	GYWRRQDRK	HLA-A24,HLA-A*2402,HLA-B*2705,HLA-Cw*0401,HLA-B*2902	55
8	NPANNAIV	HLA-A*2402,HLA-B*5301,HLA-Cw*0401	NPNNNAATV	HLA-B*5301,HLA-Cw*0401	NPNNDSAIV	HLA-Cw*0401	55

Table 3. Cont.

Conserved CD8 + T Cell Epitopes in SARS-CoV-2, SARS-CoV, and MERS-CoV							
Spike							
1	ARDLICAQK	HLA-A1,HLA-Cw*0401	ARDLICAQK	HLA-A1,HLA-Cw*0401	ARDLICAQY	HLA-A1,HLA-Cw*0401,HLA-B*2703	88
2	DLLFNKVTL	HLA-A2,HLA-A*0206,HLA-Cw*0401	DLLFNKVTL	HLA-A2,HLA-A*0206,HLA-Cw*0401	DLLFDKVTI	HLA-A3,HLA-Cw*0401	77
3	NLNESLIDL	HLA-Cw*0401	NLNESLIDL	HLA-Cw*0401	ALNESYIDL	HLA-A*0203,HLA-Cw*0401	77
4	YIKWPWYIW	HLA-A*0206,HLA-Cw*0401	YIKWPWYVW	HLA-A*0206,HLA-Cw*0401	YNKWPWYIW	HLA-A*0206,HLA-B*3501,HLA-B*5102,HLA-B*5103,HLA-B*5401,HLA-B7,HLA-Cw*0401,HLA-B35	77
5	GFIAGLIAI	HLA-B8,HLA-Cw*0401	GFIAGLIAI	HLA-B8,HLA-Cw*0401	GFIAGLVAL	HLA-B8,HLA-Cw*0401,HLA-B35	77
6	QYIKWPWYI	HLA-A*1101,HLA-A11,HLA-A24,HLA-A3,HLA-B8,HLA-Cw*0401,	QYIKWPWYV	HLA-A24,HLA-A*2402,HLA-B*5301,HLA-B8,HLA-Cw*0401	YYNKWPWYI	HLA-A*1101,HLA-A11,HLA-B8,HLA-Cw*0401,	66
7	NQKLIANQF	HLA-B27,HLA-Cw*0401,HLA-B*2703	NQKQIANQF	HLA-B14,HLA-B27,HLA-Cw*0401	NQKLIANKF	HLA-B27,HLA-Cw*0401,	66
8	VVNQNAQAL	HLA-B*3501,HLA-B*51,HLA-B*0702,HLA-Cw*0401,HLA-B35	VVNQNAQAL	HLA-B*3501,HLA-B*51,HLA-B*0702,HLA-Cw*0401,HLA-B35	AVNNNAQAL	HLA-B*51,HLA-Cw*0401,HLA-B35	66

Table 3. Cont.

Conserved CD8 + T Cell Epitopes in SARS-CoV-2, SARS-CoV, and MERS-CoV							
9	AYRFNGIGV	HLA-A*2402,HLA-Cw*0401,HLA-B*2706,HLA-A*3301	AYRFNGIGV	HLA-A*2402,HLA-Cw*0401,HLA-B*2706,HLA-A*3301	FYRLNGVGI	HLA-A*1101,HLA-A11,HLA-Cw*0401,HLA-A*3301	55
10	DFCGKGYHL	HLA-Cw*0401	DFCGKGYHL	HLA-Cw*0401	GFCGQGTHI	HLA-A*1101,HLA-A11,HLA-B8,HLA-Cw*0401,	55
11	APHGVVFLH	HLA-A*2402,HLA-A*0301,HLA-Cw*0401	APHGVVFLH	HLA-A*2402,HLA-A*0301,HLA-Cw*0401,	APNGLYFMH	HLA-B*5102,HLA-B*5103,HLA-Cw*0401	55
12	QSAPHGVVF	HLA-Cw*0401,HLA-A*3301	QAAPHGVVF	HLA-Cw*0401,	VNAPNGLYF	HLA-B*5401,HLA-B*51,HLA-B*0702,HLA-Cw*0401	44
13	VVFLHVTYV	HLA-B*51,HLA-Cw*0401	VVFLHVTYV	HLA-B*51,HLA-Cw*0401	LYFMHVGYY	HLA-A1,HLA-A*2402,HLA-A3,HLA-B*51,HLA-Cw*0401,HLA-A*3301	44
14	QMAYRFNGI	HLA-Cw*0401,HLA-A*3301,HLA-A*6801	QMAYRFNGI	HLA-Cw*0401,HLA-A*3301,HLA-A*6801	SIFYRLNGV	HLA-B*51,HLA-Cw*0401	44

Unlike E, N, and M proteins, two conserved CD8+ T cell epitopes were found for the S protein among all the HCoV, with ~44% and ~66% conservation, respectively (Table 2 and Supplementary Figure S3). These epitopes were found to be mainly restricted by HLA-C*0401, HLA-A*0206, and HLA-A*3301 (Table 2). Conservation of these two epitopes increased to ~66% and ~77% in the SARS-CoV-2, SARS-CoV, and MERS-CoV sequences (Table 3 and Supplementary Figure S4). Further analysis of epitopes in the S sequences from these three HCoV revealed 12 conserved CD8+ T cell epitopes, with ~44%, ~55%, ~66%, ~77%, and ~88% conservation (Table 3 and Supplementary Figure S4). Most of the conserved epitopes were found to be restricted by HLA-C*0401, HLA-A1, HLA-A2, HLA-A*0206, HLA-A24, HLA-A*2402, HLA-A*3301, HLA-A*6801, HLA-B*3501, HLA-B8, HLA-B27, HLA-B*2706, and HLA-B51 (Table 3).

It is important to note that all CD8+ T epitopes predicted in the E, N, M, and S sequences were found to be 100% conserved in the latest deposited 5000 SARS-CoV-2 sequences as well as in the three variants of SARS-CoV-2 (Supplementary Figure S5).

3.3. Predicted Linear B Cell Epitopes Found to Be Conserved in E, N, M, and S Protein Sequences of HCoV

E, N, M, and S consensus protein sequences of seven HCoV were used to predict linear B cell epitopes of varying length (between 10 to 20 amino acids). Even though the protein sequences showed several regions exhibiting more than 40% homology among the seven HCoV, no B cell epitopes were found in the conserved regions. Analysis of SARS-CoV-2, SARS-CoV, and MERS-CoV, however, revealed conserved B cell epitopes among these three viruses. Two B cell epitopes, with ~42% and ~50% conservation, were predicted for the E protein, while only one epitope, with ~86% conservation, was predicted for the M protein (Table 4, Supplementary Figure S4). Similarly, for the N proteins, four epitopes were found to be conserved in the N protein with ~50%, ~56%, and ~69%, while five epitopes were found to be conserved in the S protein, showing ~50%, ~56%, and ~63% conservation (Table 4, Supplementary Figure S4).

Table 4. Predicted B cell epitopes found to be conserved in E, N, M, and S protein sequences of SARS-CoV-2, SARS-CoV, and MERS-CoV. Conserved amino acids are in bold letters.

Conserved B cell epitopes in SARS-CoV-2, SARS-CoV, and MERS-CoV				
Envelope				
Nº	SARS-CoV-2	SARS-CoV	MERS-CoV	Conservation (%)
1	YSFVSEETGTLIVN	YSFVSEETGTLIVN	LPFVQERIGLFIVN	50
2	YSFVSEETGTLI	YSFVSEETGTLI	LPFVQERIGLFI	42
Membrane				
1	SMWSFNPETNILLN	SMWSFNPETNILLN	SWWSFNPETNCLLN	86
Nucleocapsid				
1	TASWFTALTQHGKEDL	TASWFTALTQHGKEEL	TVSWYTGLTQHGKVPL	69
2	YNVTQAFGRRGPEQTQ	YNVTQAFGRRGPEQTQ	FNMVQAFGLRPGDLQ	56
3	DQVILLNKHIDAYKTF	DNVILLNKHIDAYKTF	KWLELLEQNIDAYKTF	56
4	TGAIKLDDKDPNFKDQ	HGAIKLDDKDPQFKDN	SGAIKLDPKNPYNKW	50
Spike				
1	NEVAKNLNESLIDLQE	NEVAKNLNESLIDLQE	QQVVKALNESYIDLKE	63
2	ESLIDLQELGKYEQYI	ESLIDLQELGKYEQYI	ESYIDLKELGNYTYYN	63
3	CVLGQSKRVDFCGKGY	CVLGQSKRVDFCGKGY	CVKAQSKRSGFCGQGT	56
4	RDLICAQKFNGLTVLP	RDLICAQKFNGLTVLP	RDLICAQYVAGYKVLTP	50
5	EAEVQIDRLITGRLQS	EAEVQIDRLITGRLQS	EQDAQIDRLINGRLTT	50

It is again important to note that all the predicted B cell epitopes for E, N, M, and S were found to be 100% conserved in the latest deposited 5000 SARS-CoV-2 sequences as well as in the three variants (B.1.1.7, B.1.351, and P.1) of SARS-CoV-2 (Table 5, Supplementary Figure S5).

Table 5. List of immunogenic and non-toxic T and B cell epitopes from E, N, M, and S proteins. The table shows sequences of the shortlisted epitopes, their immunogenicity and toxicity score, and conservation (including level of conservation) between seven HCoV-229E and the three closely related SARS-CoV-2, SARS-CoV, and MERS-CoV as well as in three variants of SARS-CoV-2. The last column shows whether these epitopes have previously been experimentally validated or not.

SARS-CoV-2 Antigenic Non-Toxic CD8+ T and B Cell Epitopes							
T Cell Epitopes							
Envelope							
No	Peptide	Immunogenicity (Vaxijen)	Toxicity (ToxinPred)	Conservation in 7 HCoVs (%)	Conservation in SARS-CoV-2, SARS-CoV, MERS-CoV (%)	Conservation in B.1.1.7, B.1.351, P.1 (%)	Studies Experimentally Validating the Epitopes
1	SLVKPSFYV	0.414	Non-toxin	-	44	100	[58,59]
Membrane							
1	SYFIASFRL	0.4821	Non-toxin	55	66	100	[59–61]
2	AGDSGFAAY	0.9095	Non-toxin	-	44	100	[61,62]
Nucleocapsid							
1	SPRWYFYLL	0.734	Non-toxin	44	77	100	[62,63]
2	QASSRSSSR	0.8294	Non-toxin	-	66	100	[59]
3	VPINTNSSP	0.4439	Non-toxin	-	66	88	[62]
Spike							
1	DLLFNKVTLL	0.68	Non-toxin	-	77	100	[62,64,65]
2	NLNESLIDL	0.6827	Non-toxin	-	77	100	[58,59,64]
3	VVNQNAQAL	0.4749	Non-toxin	-	66	100	[64]
4	AYRFNGIGV	1.2995	Non-toxin	-	55	100	[62,64,65]
5	VVFLHVTYV	1.5122	Non-toxin	-	44	100	[62]
6	QMAYRFNGI	0.6803	Non-toxin	-	44	100	[62,65]
B cell epitopes							
Envelope							
№	Peptide	Immunogenicity (Vaxijen)	Toxicity (ToxinPred)	Conservation in 7 HCoVs (%)	Conservation in SARS-CoV-2, SARS-CoV, MERS-CoV (%)	Conservation in B.1.1.7, B.1.351, P.1 (%)	Experimentally Validated Epitope
1	YSFVSEETGLIVN	0.4532	Non-Toxin	-	50	100	[66]
2	YSFVSEETGLI	0.5014	Non-Toxin	-	42	100	[66]
Nucleocapsid							
1	TASWFTALTQHGKEDL	0.4149	Non-Toxin	-	69	100	-
2	YNVTQAFGRRGPEQTQ	0.4899	Non-Toxin	-	56	100	-
3	TGAIKLDDKDPNFKDQ	1.8438	Non-Toxin	-	50	100	-
Spike							
1	ESLIDLQELGKYEQYI	0.6105	Non-Toxin	-	63	100	-

3.4. Selection of Immunogenic Non-Toxic CD8+ T and Linear B Cell Epitopes for Design of Multi-Subunit Vaccine

In the next step, the immunogenicity and toxicity profile of all the predicted epitopes found to be conserved in the E, N, M, and S proteins of SARS-CoV-2 and other HCoVs was evaluated. Among the CD8+ T cell epitopes, one out of two epitopes for E, two out of four epitopes for M, three out of eight epitopes for N, and six out of fourteen epitopes for S were predicted to be immunogenic and non-toxic, with the highest immunogenic score (1.5) for the peptides of S, followed by M, N, and E (Table 5). From B cell epitopes, two out of two epitopes for E, three out of four epitopes for N, and one of five epitopes for the S protein

were predicted to be immunogenic and non-toxic, with the highest immunogenic score (1.8) for the peptides of N, followed by S and E (Table 5). B cell epitope for M, however, was predicted to be non-immunogenic. Therefore, the B cell epitope for M was not selected for further analysis. A total of 18 (12 CD8+ T cell and 6 linear B cell) immunogenic, non-toxic, and conserved epitopes were subsequently used to construct a multi-subunit vaccine.

3.5. Design and Assessment of the Multi-Subunit Vaccine

Physico-chemical assessment of the vaccine construct revealed the estimated half-life of the vaccine to be 30 h in mammalian reticulocytes (in vitro), >20 h in yeast (in vivo), and >10 h in *Escherichia coli* (in vivo). The instability index was predicted to be 33.06, classifying the vaccine as stable. The aliphatic index and grand average of hydropathicity (GRAVY) of the vaccine was found to be -0.250 , making the construct hydrophilic and capacitating it to interact better with the neighboring water molecules [37]. The molecular weight of the construct was found to be 34,336.60 daltons. The solubility index of the vaccine construct, based on Protein-Sol calculation, was found to be 0.34, indicating an acceptable solubility profile for the vaccine. Allergenicity and immunogenicity prediction revealed the construct to be a probable antigen (VaxiJen score 0.56) and non-allergen.

3.6. Secondary and Tertiary Structure of the Multi-epitope Vaccine:

Secondary structure assessment of the vaccine construct showed that the construct had 11 helices and 11 strands (Figure 3). The 3D structure of the vaccine was constructed using both the Phyre2 and I-TASSER tools. The structure constructed using I-TASSER had an overall low-quality score after refinement, and on a Ramachandran plot, only 88% residues were in the favorable region (data not shown). To the contrary, the 3D structure predicted by the Phyre2 tool after refinement (Figure 4A) exhibited a high confidence score, ranging from 84–96% for most regions, and 66–74% for some regions. The Ramachandran analysis of the structure showed that the predicted structure had 94.3% of the residues in a highly preferred region, 4.5% in a preferred region, and only 1.1% residues in a non-preferred/disallowed region (Figure 4B). The PROSA score of the structure predicted by Phyre2 was -2.94 , which lies in the acceptable range (Figure 4C). Overall, the assessment of the different protein quality parameters indicates the structure predicted by Phyre2 to be of high quality.

3.7. Molecular Docking Analysis of the Vaccine–TLR4 Complex

To reveal the binding modes and energies of the developed vaccine with human TLR4, the molecular docking technique was utilized. All possible vaccine–TLR4 solutions (exactly 1000 solutions) were first predicted using the PatchDock server [50]. The predicted vaccine–TLR4 complex solutions were then refined, and the global binding energies were rescored using the FireDock server [51,52]. The 3D representation of the predicted vaccine–TLR4 complex structure is depicted in Figure 5, and the estimated global binding energy and energy contributions for the top-ranked vaccine–TLR4 complex are listed in Table 6. Interactions between the vaccine–TLR4 complex were resolved using LigPlot+ software.

The global binding energy estimates suggest that our vaccine construct formed a stable complex with TLR4, with global binding energy of -47.76 kcal/mol (Table 6 and Figure 5A). Assessment of the energy contributions revealed that attractive van der Waals (vdW) was the dominant force in vaccine–TLR4 binding energy, with a value of -36.33 kcal/mol (Table 6). Analysis of the protein–protein interaction suggested that the construct formed multiple interactions with TLR4, especially with several residues of β -defensin (adjuvant used in our construct, and a known agonist of TLR4 [46,47] were also found to interact with TLR4 (Figure 5B).

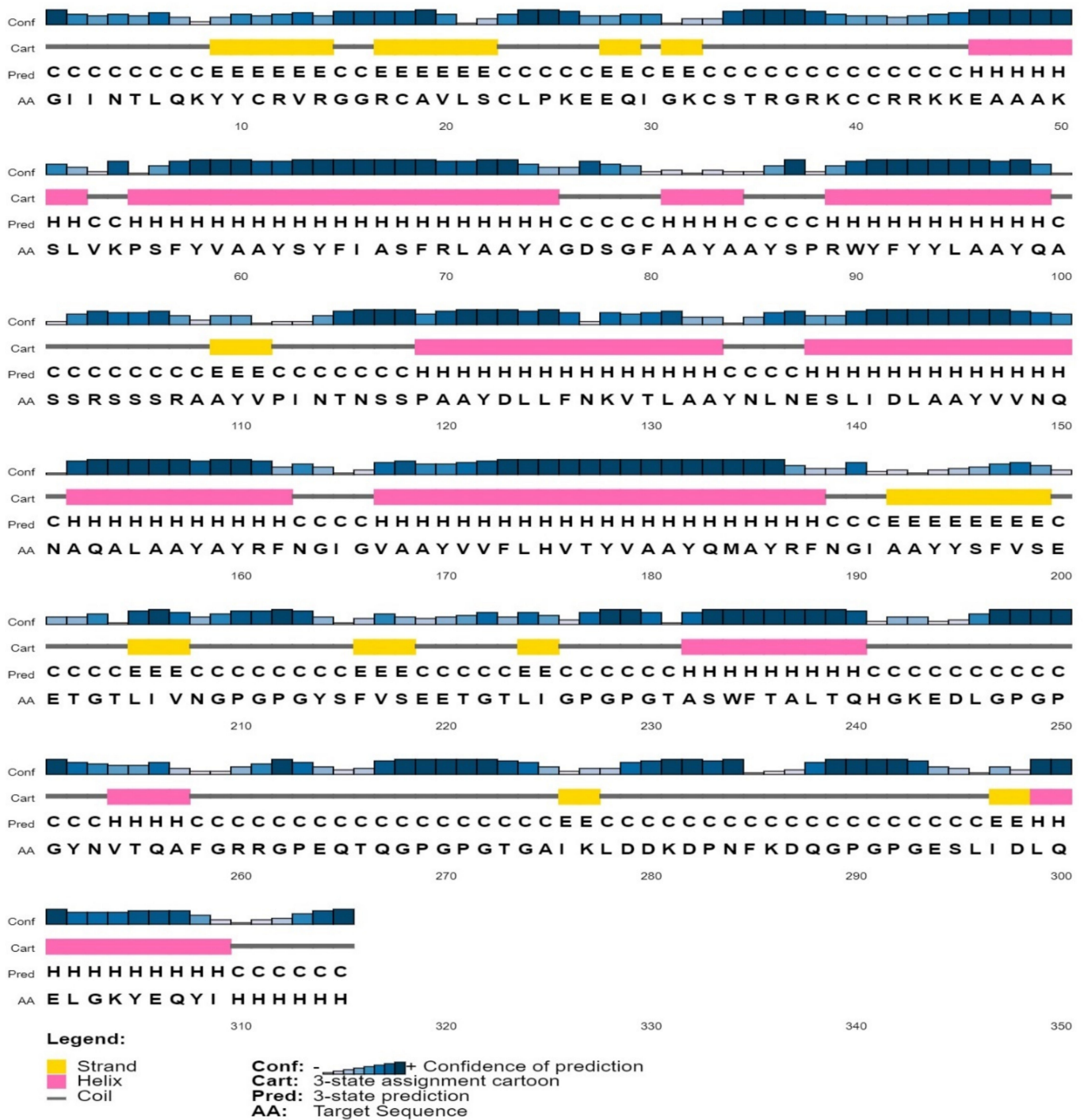


Figure 3. Secondary structure prediction of the vaccine construct. The color-coding of different regions is explained in the legend within the figure. For secondary structure prediction, C, H, and E symbols indicate the coil, helix, and strand regions, respectively.

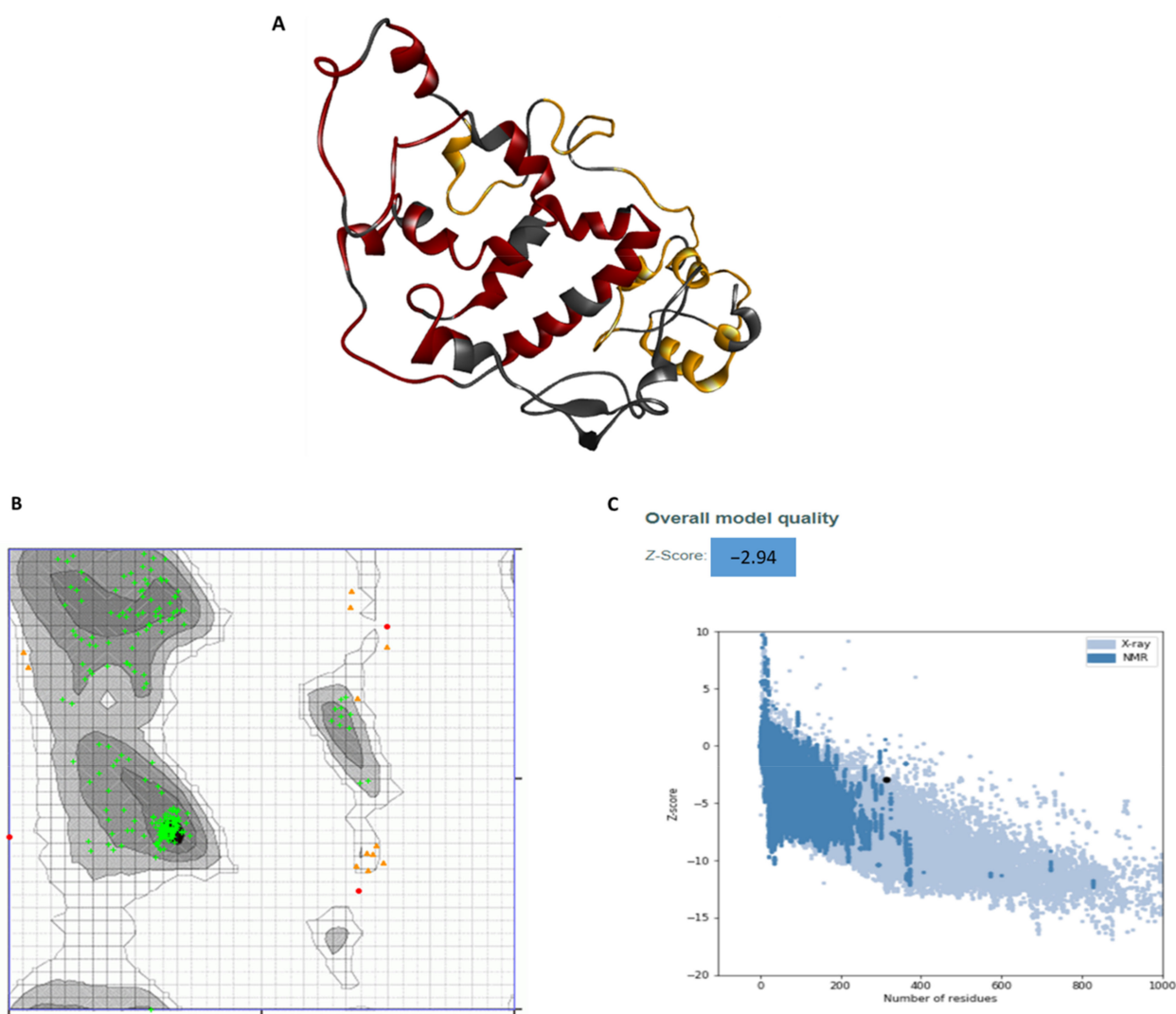


Figure 4. Predicted 3D structure of the vaccine and its quality assessment. (A) Predicted 3D structure of the vaccine (black) showing T cell epitopes in red, while B cell epitopes in orange, (B) Ramachandran plot for the 3D structure of the vaccine showing 97.8% of the residues in the favorable region, (C) ProSA z-score for the 3D structure of the vaccine, determined using NMR spectroscopy (dark blue) and X-ray crystallography (light blue).

Table 6. Docking results of the developed vaccine with human toll-like receptor 4 (TLR4). The table shows global binding energy as well as different forces contributing to the global energy.

Complex	Global Binding Energy	Energy Contributions			
		Attractive Van Der Waals (vdW)	Repulsive Van Der Waals (vdW)	Atomic Contact Energy (ACE)	Hydrogen Bond (HB)
Vaccine-TLR4	−47.76	−36.33	16.90	15.29	−0.42

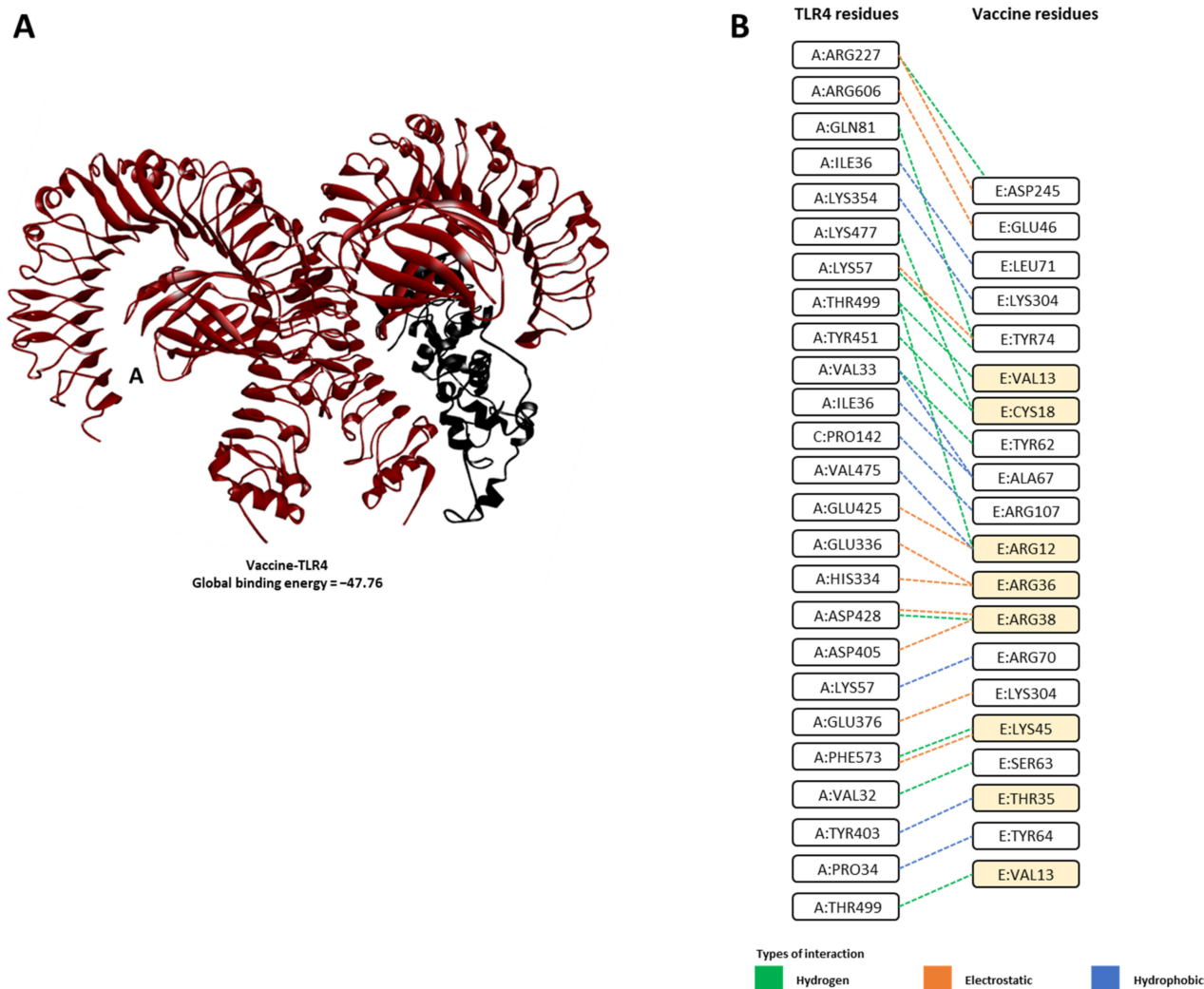


Figure 5. Molecular docking analysis between vaccine construct and TLR4. (A) Docked complex of the designed vaccine with TLR4, (B) amino acid interaction between vaccine-TLR4 (vaccine construct: chain E; TLR4: chain A; C: Lymphocyte antigen 96), resolved using LigPlot+ (DIMPLLOT). Residues that are part of β -defensin found to interact with TLR4 are highlighted in orange.

3.8. In Silico Simulation of the Immune Response against the Vaccine Construct

We performed in silico immune simulations to examine the potential of our vaccine construct to elicit potent and durable B (including antibody response) and T cell responses. Results from proteasomal processing showed that the proteasomal sites in the vaccine are strategically located to release all 9-mer T cell epitopes and variable-length B cell epitopes (Supplementary Figure S2). Results from the immune simulation showed that the vaccine construct was predicted to elicit significant titers of IgG and IgM (Figure 6A), significant effector and memory B-cell response (Figure 6B) as well as cytotoxic T cell and interferon-gamma response (Figure 6C). Overall, the vaccine construct was predicted to be stable, non-toxic and highly immunogenic, and capable of mounting a sustained protective immune response.

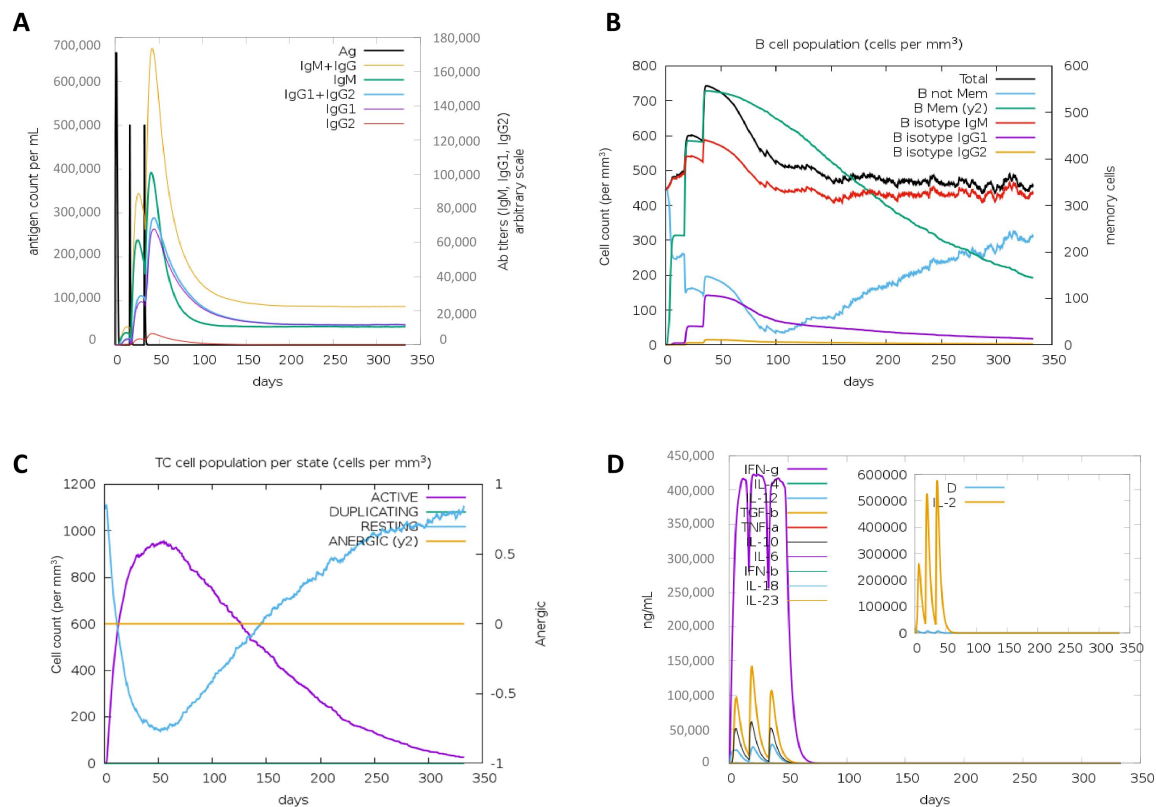


Figure 6. Simulated immune response after administration of the vaccine construct: The graphs show stimulation of (A) IgG, IgM, and total antibody response, (B) Effector and memory B cell response, (C) Active and anergic CD8+ T cell response, and (D) IFN-gamma and other cytokine responses after vaccine administration; (D) (insert): plot shows levels of danger signal (positive or negative signals immune systems generate in response to ‘danger’ such as pathogens (Ramadan et al., 2017)) together with leukocyte growth factor IL-2 over days.

4. Discussion

Using immunoinformatic tools, we have designed, constructed, evaluated, and assessed the performance of a non-allergenic and antigenic chimeric multi-epitope vaccine that combines T and B cell epitopes from conserved regions of the E, M, N, and S proteins of seven human coronaviruses. The vaccine shows potential to elicit a strong response against SARS-CoV-2 while also providing immune protection against emerging SARS-CoV-2 variants and other HCoVs.

All of the currently available vaccines target the SARS-CoV-2 S protein, which is the most antigenic, highly variable, and prone to mutations [19,20]. Due to the high substitution/mutation rate and rapid transmissions, new variants of SARS-CoV-2 have been continuously emerging globally, posing threats to circumvent the efficacy of current vaccines, possibly leading to a novel-SARS-CoV-2-variant-associated surge in transmissions [15]. In fact, there have already been reports of the current vaccines losing effectiveness against the emerging SARS-CoV-2 variants [21,22]. Variants of high concern, B.1.1.7 (VOC-202012/01) identified in the U.K.; 501Y.V2 (B.1.351) identified in the South Africa; and P.1 (B.1.1.28.1) identified in Brazil, all have more than 20 mutations in the S protein, raising the concern that both natural infection-derived immunity as well as vaccine-derived immunity may fail to confer protection against a SARS-CoV-2 variant (re)infection [15,21]. This hypothesis has been partly supported by a recent study showing that the 501Y.V2 (B.1.351) variant presents a complete immune escape from natural immunity [67]. Nonetheless, efforts to design new vaccines or to improve the existing vaccines that can confer immunity against SARS-CoV-2 and its variants are underway. In this regard, numerous in silico studies have also been conducted to design multi-epitope vaccine candidates against SARS-CoV-2.

However, most of these studies focused on epitopes from SARS-CoV-2 S protein as the antigenic targets [37,68–71], while others included the N and M proteins [72,73].

The primary structure of the protein can provide information regarding mutations and the conservation of residues, while the secondary structure, depicting the spatial orientation, such as β -folds and α -helices, can give insights into how the protein will fold into its final form. Finally, the tertiary structure depicts the three-dimensional orientation of the protein in space [74]. The nature of the amino acids and protein folding can affect the physico-chemical properties of the protein, such as solubility, stability and interaction with other molecules, etc. [74]. The physico-chemical analyses revealed the molecular weight of our vaccine, 34.34 kDa, indicating that it is an ideal vaccine candidate, as proteins with a molecular weight less than 100 kDa are considered to be efficient vaccine candidates [70]. The instability index of our vaccine construct was shown to be 33.06, indicating that our construct is stable, as proteins with an index greater than 40 are generally considered unstable [38]. The vaccine construct was also found to be thermo-stable, with an aliphatic index of 72.00. A high aliphatic index is an indication of a protein to be thermo-stable over a wide temperature range [75]. The GRAVY index was predicted to be -0.250 , reflecting high solubility of the vaccine and its effective interaction with water molecules [76]. The half-life of the vaccine was found to be 30 h in mammalian reticulocytes (in vitro), >20 h in yeast (in vivo), and >10 h in *Escherichia coli* (in vivo). Vaccine candidates with similar a half-life have been found to induce strong humoral and cellular responses in mice [77]. Finally, a refined and validated vaccine model showed that the designed vaccine is of high quality, with $>97\%$ of residues in the favored region [78].

In our study, we analyzed all the structural proteins, E, N, M, and S, not only in SARS-CoV-2 and its known variants, but also from six other human coronaviruses, and selected highly conserved epitopes with the aim of designing an effective multi-subunit vaccine that can be effective against SARS-CoV-2 and its variants and may also confer cross-immunoprotection against other HCoV. Yazdani et al. [79] conducted a similar study using the E, N, M, and S proteins, however, this work investigated the conservation of peptides for only SARS-CoV-2 and SARS-CoV. In addition, the final epitopes selected in their study presented a lower antigenicity score (~ 0.7), while the epitopes we used in our design exhibited a high immunogenicity score of 1.5 for CD8⁺ T epitopes residing in S protein, and of 1.8 for linear B cell epitopes residing in the N protein. As our multi-subunit vaccine contains peptides from multiple proteins which are 100% conserved in the B.1.1.7, 501Y.V2, and P.1 variants SARS-CoV-2, mutations that potentially reduce immunogenicity are not likely to occur. The existence of cross-immune reactivity has been documented in several studies [80–84]. It has been reported that the antigenic domains of the S and N proteins are highly cross-reactive across coronaviruses [80,82,83], and it has been shown that T cell immunity previously induced by circulating human alpha- and beta-HCoVs in young adults is protective against severe clinical outcomes [81].

An ideal vaccine candidate, on the one hand, should be immunogenic/antigenic, and should be able to elicit targeted humoral and cellular immune responses, while, on the other hand, should also be non-toxic and non-allergenic [85], hence providing protection against adverse effects. Vaccines using a complete agent (virus, for example) or large proteins carry an increased chance of allergenic response due to the unnecessary antigenic load introduced in the vaccine design [86]. This drawback is avoided in our construct, as we used short immunogenic peptides that were predicted to be non-allergenic and non-toxic. Moreover, the predicted conserved T and B cell epitopes used for our multi-subunit vaccine can elicit cellular and humoral immune responses simultaneously, further increasing the chance of triggering an efficient overall immune response. Additionally, our vaccine was constructed with β -defensin as an adjuvant at the N-terminal, which allows for a long-lasting immune response. β -defensin has been shown to be an efficient adjuvant capable of enhancing the immunogenicity of candidate vaccines against viruses, including MERS-CoV [14,87].

In the immune simulation analyses, these properties were predicted for our vaccine construct as well, showing that the vaccine may be able to elicit a sustained IgG, IgM, B- and cytotoxic T-response, and interferon-gamma response for well over 300 days. The activation/release of interferon-gamma is a hallmark of antiviral activity by T cells, as this cytokine is capable of killing the virus-infected cells [88]. Similarly, the analysis of the immunological response in individuals enrolled in phase 1/2 trial of the ChAdOx1 nCoV-19 vaccine showed that a single dose of the vaccine in adults was able to induce a potent immune response, primarily composed of neutralization antibodies of IgG1 and IgG3 subclasses, TH-1 response with interferon- γ , and tumor necrosis factor- α cytokine secretion by CD4+ T cells as well as mono- and poly-functional and cytotoxic CD8+ T cells up to 8 weeks after vaccination [89]. Similar results were observed for the Pfizer/BioNTech and Moderna vaccines, where robust T cell response against spike protein of SARS-CoV-2 was observed in vaccinated individuals, which was more profound than the response observed in convalescent patients [90]. Moreover, the CD8+ T cell epitopes in our vaccine design were linked with the AYY linker that provides binding sites for the transporter associated with antigen processing (TAP) to facilitate epitope presentation [91], while B cell epitopes were linked with the glycine rich GPGPG linker, which gives the vaccine structural flexibility, prevents junctional epitope formation, and helps immune processing and epitope presentation [91].

Lastly, in molecular docking studies, using TLR4 showed that our vaccine construct, and especially several of β -defensin residues, a known TLR4 agonist [46,47] used as an adjuvant in our vaccine, can form stable complexes with the TLR4. TLR-4 is capable of recognizing various exogenous and endogenous ligands such as viral protein, lipopolysaccharide, heat shock proteins, β -defensin, etc., which leads to dendritic cell maturation and up-regulation of costimulatory molecules, resulting in a type 1 polarized adaptive immune response [46,47]. Recently, Zhao et al. showed that SARS-CoV-2 is able to activate TLR4 to activate an anti-bacterial like immune response [92]. Similarly, other in silico studies have shown that the ability of a vaccine construct to form stable complexes with these TLRs may lead to the induction of antiviral immunity [37,93].

5. Conclusions

Multi-epitope vaccines designed using immune-informatics have been widely studied against viruses as well as against cancer [94–97]. Some of these vaccines have shown promising results in vivo, while some entered clinical trials [98–102]. Analyzing a vaccine construct in silico before proceeding to in vitro evaluations and clinical trials is an efficient way to determine the vaccine's potential with confidence, economizing on time, effort, and money. Here we present in silico analyses of a multi-subunit vaccine design that, from multiple perspectives, shows potential to elicit a strong immune-protective response against SARS-CoV-2 and its emerging variants, while carrying minimal risk for causing adverse effects.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/vaccines9070702/s1>, Figure S1: Multisubunit vaccine construct, Figure S2: Phylogenetic analysis between human coronaviruses, Figure S3: Conservation of envelope, membrane, nucleocapsid and spike protein sequences in seven HCoV, Figure S4: Conservation of envelope, membrane, nucleocapsid, and spike protein sequences in SARS-CoV-2, SARS-CoV and MERS-CoV, Figure S5: Conservation of envelope, membrane, nucleocapsid, and spike protein sequence in 5000 SARS-CoV-2 sequences and in B.1.1.7, B.1.351, and P.1 variants of SARS-CoV-2, Table S1: Consensus sequences used to predict the epitopes for vaccine design, Table S2: List of all epitopes predicted with CTLpred.

Author Contributions: B.A.: methodology, analysis, and manuscript writing, M.A.A.I: methodology, analysis, and manuscript writing, Z.M.: analysis and manuscript writing, S.H.A.: conceptualization, methodology, analysis, and manuscript writing, S.A.: conceptualization, methodology, manuscript writing, and supervision. All authors have read and agreed to the published version of the manuscript.

Funding: Not applicable.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All datasets presented in this study are included in the article/Supplementary Materials.

Conflicts of Interest: Authors declare no conflicts of interests.

References

1. Zhou, P.; Yang, X.L.; Wang, X.G.; Hu, B.; Zhang, L.; Zhang, W.; Si, H.R.; Zhu, Y.; Li, B.; Huang, C.L.; et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **2020**, *579*, 270–273; Addendum in **2020**, *588*, E6. [[CrossRef](#)] [[PubMed](#)]
2. Wu, Z.; McGoogan, J.M. Characteristics of and Important Lessons From the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72314 Cases From the Chinese Center for Disease Control and Prevention. *JAMA* **2020**, *323*, 1239–1242. [[CrossRef](#)]
3. Corman, V.M.; Muth, D.; Niemeyer, D.; Drosten, C. Hosts and Sources of Endemic Human Coronaviruses. *Adv. Virus Res.* **2018**, *100*, 163–188. [[CrossRef](#)] [[PubMed](#)]
4. De Groot, R.J.; Baker, S.C.; Baric, R.S.; Brown, C.S.; Drosten, C.; Enjuanes, L.; Fouchier, R.A.; Galiano, M.; Gorbalenya, A.E.; Memish, Z.A.; et al. Middle East respiratory syndrome coronavirus (MERS-CoV): Announcement of the Coronavirus Study Group. *J. Virol.* **2013**, *87*, 7790–7792. [[CrossRef](#)]
5. Perlman, S.; Netland, J. Coronaviruses post-SARS: Update on replication and pathogenesis. *Nat. Rev. Microbiol.* **2009**, *7*, 439–450. [[CrossRef](#)]
6. Fouchier, R.A.; Hartwig, N.G.; Bestebroer, T.M.; Niemeyer, B.; de Jong, J.C.; Simon, J.H.; Osterhaus, A.D. A previously undescribed coronavirus associated with respiratory disease in humans. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 6212–6216. [[CrossRef](#)]
7. Lim, Y.X.; Ng, Y.L.; Tam, J.P.; Liu, D.X. Human Coronaviruses: A Review of Virus-Host Interactions. *Diseases* **2016**, *4*, 26. [[CrossRef](#)]
8. Woo, P.C.; Lau, S.K.; Chu, C.M.; Chan, K.H.; Tsoi, H.W.; Huang, Y.; Wong, B.H.; Poon, R.W.; Cai, J.J.; Luk, W.K.; et al. Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia. *J. Virol.* **2005**, *79*, 884–895. [[CrossRef](#)]
9. Hu, B.; Guo, H.; Zhou, P.; Shi, Z.L. Characteristics of SARS-CoV-2 and COVID-19. *Nat. Rev. Microbiol.* **2021**, *19*, 141–154. [[CrossRef](#)]
10. Hoffmann, M.; Kleine-Weber, H.; Schroeder, S.; Kruger, N.; Herrler, T.; Erichsen, S.; Schiergens, T.S.; Herrler, G.; Wu, N.H.; Nitsche, A.; et al. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* **2020**, *181*, 271–280e278. [[CrossRef](#)]
11. Letko, M.; Munster, V. Functional assessment of cell entry and receptor usage for lineage B beta-coronaviruses, including 2019-nCoV. *bioRxiv* **2020**. [[CrossRef](#)]
12. Habibzadeh, P.; Stoneman, E.K. The Novel Coronavirus: A Bird's Eye View. *Int. J. Occup. Environ. Med.* **2020**, *11*, 65–71. [[CrossRef](#)]
13. Shang, W.; Yang, Y.; Rao, Y.; Rao, X. The outbreak of SARS-CoV-2 pneumonia calls for viral vaccines. *NPJ Vaccines* **2020**, *5*, 18. [[CrossRef](#)]
14. Kim, J.H.; Marks, F.; Clemens, J.D. Looking beyond COVID-19 vaccine phase 3 trials. *Nat. Med.* **2021**, *27*, 205–211. [[CrossRef](#)]
15. Karim, S.S.A. Vaccines and SARS-CoV-2 variants: The urgent need for a correlate of protection. *Lancet* **2021**, *397*, 1263–1264. [[CrossRef](#)]
16. Creech, C.B.; Walker, S.C.; Samuels, R.J. SARS-CoV-2 Vaccines. *JAMA* **2021**, *325*, 1318–1320. [[CrossRef](#)]
17. Jones, I.; Roy, P. Sputnik V COVID-19 vaccine candidate appears safe and effective. *Lancet* **2021**, *397*, 642–643. [[CrossRef](#)]
18. Logunov, D.Y.; Dolzhikova, I.V.; Shcheblyakov, D.V.; Tukhvatulin, A.I.; Zubkova, O.V.; Dzharullaeva, A.S.; Kovyrshina, A.V.; Lubenets, N.L.; Grousova, D.M.; Erokhova, A.S.; et al. Safety and efficacy of an rAd26 and rAd5 vector-based heterologous prime-boost COVID-19 vaccine: An interim analysis of a randomised controlled phase 3 trial in Russia. *Lancet* **2021**, *397*, 671–681. [[CrossRef](#)]
19. Duan, L.; Zheng, Q.; Zhang, H.; Niu, Y.; Lou, Y.; Wang, H. The SARS-CoV-2 Spike Glycoprotein Biosynthesis, Structure, Function, and Antigenicity: Implications for the Design of Spike-Based Vaccine Immunogens. *Front Immunol.* **2020**, *11*, 576622. [[CrossRef](#)]
20. Li, Q.; Wu, J.; Nie, J.; Zhang, L.; Hao, H.; Liu, S.; Zhao, C.; Zhang, Q.; Liu, H.; Nie, L.; et al. The Impact of Mutations in SARS-CoV-2 Spike on Viral Infectivity and Antigenicity. *Cell* **2020**, *182*, 1284–1294e1289. [[CrossRef](#)]
21. Emary, K.R.W.; Golubchik, T.; Aley, P.K.; Ariani, C.V.; Angus, B.; Bibi, S.; Blane, B.; Bonsall, D.; Cicconi, P.; Charlton, S.; et al. Efficacy of ChAdOx1 nCoV-19 (AZD1222) vaccine against SARS-CoV-2 variant of concern 202012/01 (B.1.1.7): An exploratory analysis of a randomised controlled trial. *Lancet* **2021**, *397*, 1351–1362. [[CrossRef](#)]

22. Madhi, S.A.; Baillie, V.; Cutland, C.L.; Voysey, M.; Koen, A.L.; Fairlie, L.; Padayachee, S.D.; Dheda, K.; Barnabas, S.L.; Bhorat, Q.E.; et al. Efficacy of the ChAdOx1 nCoV-19 Covid-19 Vaccine against the B.1.351 Variant. *N. Engl. J. Med.* **2021**. [[CrossRef](#)]
23. Janice Oh, H.L.; Ken-En Gan, S.; Bertoletti, A.; Tan, Y.J. Understanding the T cell immune response in SARS coronavirus infection. *Emerg. Microbes Infect.* **2012**, *1*, e23. [[CrossRef](#)]
24. Pinotti, F.; Wikramaratna, P.S.; Obolski, U.; Paton, R.S.; Damineli, D.S.C.; Alcantara, L.C.J.; Giovanetti, M.; Gupta, S.; Lourenco, J. Potential impact of individual exposure histories to endemic human coronaviruses on age-dependent severity of COVID-19. *BMC Med.* **2021**, *19*, 19. [[CrossRef](#)]
25. Trifinopoulos, J.; Nguyen, L.-T.; von Haeseler, A.; Minh, B.Q. W-IQ-TREE: A fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Res.* **2016**, *44*, W232–W235. [[CrossRef](#)]
26. Katoh, K.; Standley, D.M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780. [[CrossRef](#)]
27. Robert, X.; Gouet, P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **2014**, *42*, W320–W324. [[CrossRef](#)]
28. Waterhouse, A.M.; Procter, J.B.; Martin, D.M.A.; Clamp, M.; Barton, G.J. Jalview Version 2—A multiple sequence alignment editor and analysis workbench. *Bioinformatics* **2009**, *25*, 1189–1191. [[CrossRef](#)]
29. Hall, T.A. BioEdit: A User-Friendly Biological Sequence Alignment Editor and Analysis Program for Windows 95/98/NT. *Nucleic Acids Symposium Series* **1999**, *41*, 95–98.
30. Larsson, A. AliView: A fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics* **2014**, *30*, 3276–3278. [[CrossRef](#)]
31. Bhasin, M.; Raghava, G.P. Prediction of CTL epitopes using QM, SVM and ANN techniques. *Vaccine* **2004**, *22*, 3195–3204. [[CrossRef](#)]
32. Bhasin, M.; Raghava, G.P. A hybrid approach for predicting promiscuous MHC class I restricted T cell epitopes. *J. Biosci.* **2007**, *32*, 31–42. [[CrossRef](#)]
33. Saha, S.; Raghava, G.P. Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. *Proteins* **2006**, *65*, 40–48. [[CrossRef](#)]
34. Vita, R.; Mahajan, S.; Overton, J.A.; Dhanda, S.K.; Martini, S.; Cantrell, J.R.; Wheeler, D.K.; Sette, A.; Peters, B. The Immune Epitope Database (IEDB): 2018 update. *Nucleic Acids Res* **2019**, *47*, D339–D343. [[CrossRef](#)]
35. Doytchinova, I.A.; Flower, D.R. VaxiJen: A server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinformatics* **2007**, *8*, 4. [[CrossRef](#)]
36. Gupta, S.; Kapoor, P.; Chaudhary, K.; Gautam, A.; Kumar, R.; Open Source Drug Discovery, C.; Raghava, G.P. In silico approach for predicting toxicity of peptides and proteins. *PLoS ONE* **2013**, *8*, e73957. [[CrossRef](#)]
37. Abraham Peele, K.; Srihansa, T.; Krupanidhi, S.; Ayyagari, V.S.; Venkateswarulu, T.C. Design of multi-epitope vaccine candidate against SARS-CoV-2: A in-silico study. *J. Biomol. Struct. Dyn.* **2020**, 1–9. [[CrossRef](#)]
38. Ikram, A.; Zaheer, T.; Awan, F.M.; Obaid, A.; Naz, A.; Hanif, R.; Paracha, R.Z.; Ali, A.; Naveed, A.K.; Janjua, H.A. Exploring NS3/4A, NS5A and NS5B proteins to design conserved subunit multi-epitope vaccine against HCV utilizing immunoinformatics approaches. *Sci. Rep.* **2018**, *8*, 16107. [[CrossRef](#)]
39. Dimitrov, I.; Bangov, I.; Flower, D.R.; Doytchinova, I. AllerTOP v.2—A server for in silico prediction of allergens. *J. Mol. Model.* **2014**, *20*, 2278. [[CrossRef](#)]
40. McGuffin, L.J.; Bryson, K.; Jones, D.T. The PSIPRED protein structure prediction server. *Bioinformatics* **2000**, *16*, 404–405. [[CrossRef](#)]
41. 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](#)]
42. Kelley, L.A.; Mezulis, S.; Yates, C.M.; Wass, M.N.; Sternberg, M.J. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **2015**, *10*, 845–858. [[CrossRef](#)]
43. Heo, L.; Park, H.; Seok, C. GalaxyRefine: Protein structure refinement driven by side-chain repacking. *Nucleic Acids Res.* **2013**, *41*, W384–W388. [[CrossRef](#)]
44. Wiederstein, M.; Sippl, M.J. ProSA-web: Interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res.* **2007**, *35*, W407–W410. [[CrossRef](#)]
45. Hebditch, M.; Carballo-Amador, M.A.; Charonis, S.; Curtis, R.; Warwicker, J. Protein-Sol: A web tool for predicting protein solubility from sequence. *Bioinformatics* **2017**, *33*, 3098–3100. [[CrossRef](#)]
46. Biragyn, A.; Ruffini, P.A.; Leifer, C.A.; Klyushnenkova, E.; Shakhov, A.; Chertov, O.; Shirakawa, A.K.; Farber, J.M.; Segal, D.M.; Oppenheim, J.J.; et al. Toll-like receptor 4-dependent activation of dendritic cells by beta-defensin 2. *Science* **2002**, *298*, 1025–1029. [[CrossRef](#)]
47. Brubaker, S.W.; Bonham, K.S.; Zanoni, I.; Kagan, J.C. Innate immune pattern recognition: A cell biological perspective. *Annu Rev. Immunol* **2015**, *33*, 257–290. [[CrossRef](#)]
48. Ohto, U.; Yamakawa, N.; Akashi-Takamura, S.; Miyake, K.; Shimizu, T. Structural Analyses of Human Toll-like Receptor 4 Polymorphisms D299G and T399I*. *Journal of Biological Chemistry* **2012**, *287*, 40611–40617. [[CrossRef](#)]
49. Gordon, J.C.; Myers, J.B.; Folta, T.; Shoja, V.; Heath, L.S.; Onufriev, A. H++: A server for estimating pKas and adding missing hydrogens to macromolecules. *Nucleic Acids Res.* **2005**, *33*, W368–W371. [[CrossRef](#)]

50. Schneidman-Duhovny, D.; Inbar, Y.; Nussinov, R.; Wolfson, H.J. PatchDock and SymmDock: Servers for rigid and symmetric docking. *Nucleic Acids Res.* **2005**, *33*, W363–W367. [[CrossRef](#)]
51. Andrusier, N.; Nussinov, R.; Wolfson, H.J. FireDock: Fast interaction refinement in molecular docking. *Proteins* **2007**, *69*, 139–159. [[CrossRef](#)]
52. Mashiach, E.; Schneidman-Duhovny, D.; Andrusier, N.; Nussinov, R.; Wolfson, H.J. FireDock: A web server for fast interaction refinement in molecular docking. *Nucleic Acids Res.* **2008**, *36*, W229–W232. [[CrossRef](#)]
53. Laskowski, R.A.; Swindells, M.B. LigPlot+: Multiple ligand-protein interaction diagrams for drug discovery. *J. Chem. Inf. Model.* **2011**, *51*, 2778–2786. [[CrossRef](#)]
54. Rapin, N.; Lund, O.; Castiglione, F. Immune system simulation online. *Bioinformatics* **2011**, *27*, 2013–2014. [[CrossRef](#)]
55. 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](#)]
56. Kesmir, C.; Nussbaum, A.K.; Schild, H.; Detours, V.; Brunak, S. Prediction of proteasome cleavage motifs by neural networks. *Protein Eng.* **2002**, *15*, 287–296. [[CrossRef](#)]
57. Nielsen, M.; Lundegaard, C.; Lund, O.; Kesmir, C. The role of the proteasome in generating cytotoxic T-cell epitopes: Insights obtained from improved predictions of proteasomal cleavage. *Immunogenetics* **2005**, *57*, 33–41. [[CrossRef](#)]
58. Kared, H.; Redd, A.D.; Bloch, E.M.; Bonny, T.S.; Sumatoh, H.; Kairi, F.; Carbajo, D.; Abel, B.; Newell, E.W.; Bettinotti, M.P.; et al. SARS-CoV-2-specific CD8+ T cell responses in convalescent COVID-19 individuals. *J. Clin. Investig.* **2021**, *131*. [[CrossRef](#)]
59. Snyder, T.M.; Gittelman, R.M.; Klinger, M.; May, D.H.; Osborne, E.J.; Taniguchi, R.; Zahid, H.J.; Kaplan, I.M.; Dines, J.N.; Noakes, M.N.; et al. Magnitude and Dynamics of the T-Cell Response to SARS-CoV-2 Infection at Both Individual and Population Levels. *medRxiv* **2020**. [[CrossRef](#)]
60. Bilich, T.; Nelde, A.; Heitmann, J.S.; Maringer, Y.; Roerden, M.; Bauer, J.; Rieth, J.; Wacker, M.; Peter, A.; Horber, S.; et al. T cell and antibody kinetics delineate SARS-CoV-2 peptides mediating long-term immune responses in COVID-19 convalescent individuals. *Sci. Transl. Med.* **2021**, *13*. [[CrossRef](#)]
61. Schwarz, T.; Heiss, K.; Mahendran, Y.; Casilag, F.; Kurth, F.; Sander, L.E.; Wendtner, C.M.; Hoechstetter, M.A.; Muller, M.A.; Sekul, R.; et al. SARS-CoV-2 Proteome-Wide Analysis Revealed Significant Epitope Signatures in COVID-19 Patients. *Front. Immunol.* **2021**, *12*, 629185. [[CrossRef](#)]
62. Tarke, A.; Sidney, J.; Kidd, C.K.; Dan, J.M.; Ramirez, S.I.; Yu, E.D.; Mateus, J.; da Silva Antunes, R.; Moore, E.; Rubiro, P.; et al. Comprehensive analysis of T cell immunodominance and immunoprevalence of SARS-CoV-2 epitopes in COVID-19 cases. *Cell Rep. Med.* **2021**, *2*, 100204. [[CrossRef](#)]
63. Ferretti, A.P.; Kula, T.; Wang, Y.; Nguyen, D.M.V.; Weinheimer, A.; Dunlap, G.S.; Xu, Q.; Nabili, N.; Perullo, C.R.; Cristofaro, A.W.; et al. Unbiased Screens Show CD8(+) T Cells of COVID-19 Patients Recognize Shared Epitopes in SARS-CoV-2 that Largely Reside outside the Spike Protein. *Immunity* **2020**, *53*, 1095–1107e1093. [[CrossRef](#)]
64. Li, Y.; Ma, M.L.; Lei, Q.; Wang, F.; Hong, W.; Lai, D.Y.; Hou, H.; Xu, Z.W.; Zhang, B.; Chen, H.; et al. Linear epitope landscape of the SARS-CoV-2 Spike protein constructed from 1,051 COVID-19 patients. *Cell Rep.* **2021**, *34*, 108915. [[CrossRef](#)]
65. Mateus, J.; Grifoni, A.; Tarke, A.; Sidney, J.; Ramirez, S.I.; Dan, J.M.; Burger, Z.C.; Rawlings, S.A.; Smith, D.M.; Phillips, E.; et al. Selective and cross-reactive SARS-CoV-2 T cell epitopes in unexposed humans. *Science* **2020**, *370*, 89–94. [[CrossRef](#)]
66. Lu, S.; Xie, X.X.; Zhao, L.; Wang, B.; Zhu, J.; Yang, T.R.; Yang, G.W.; Ji, M.; Lv, C.P.; Xue, J.; et al. The immunodominant and neutralization linear epitopes for SARS-CoV-2. *Cell Rep.* **2021**, *34*, 108666. [[CrossRef](#)]
67. Wibmer, C.K.; Ayres, F.; Hermanus, T.; Madzivhandila, M.; Kgagudi, P.; Oosthuysen, B.; Lambson, B.E.; de Oliveira, T.; Vermeulen, M.; van der Berg, K.; et al. SARS-CoV-2 501Y.V2 escapes neutralization by South African COVID-19 donor plasma. *Nat. Med.* **2021**, *27*, 622–625. [[CrossRef](#)]
68. Sanami, S.; Zandi, M.; Pourhossein, B.; Mobini, G.R.; Safaei, M.; Abed, A.; Arvejeh, P.M.; Chermahini, F.A.; Alizadeh, M. Design of a multi-epitope vaccine against SARS-CoV-2 using immunoinformatics approach. *Int. J. Biol. Macromol.* **2020**, *164*, 871–883. [[CrossRef](#)]
69. Kar, T.; Narsaria, U.; Basak, S.; Deb, D.; Castiglione, F.; Mueller, D.M.; Srivastava, A.P. A candidate multi-epitope vaccine against SARS-CoV-2. *Sci. Rep.* **2020**, *10*, 10895. [[CrossRef](#)]
70. Naz, A.; Awan, F.M.; Obaid, A.; Muhammad, S.A.; Paracha, R.Z.; Ahmad, J.; Ali, A. Identification of putative vaccine candidates against *Helicobacter pylori* exploiting exoproteome and secretome: A reverse vaccinology based approach. *Infect. Genet. Evol.* **2015**, *32*, 280–291. [[CrossRef](#)]
71. Martin, W.R.; Cheng, F. A rational design of a multi-epitope vaccine against SARS-CoV-2 which accounts for the glycan shield of the spike glycoprotein. *J. Biomol. Struct. Dyn.* **2021**, 1–15. [[CrossRef](#)]
72. Sarkar, B.; Ullah, M.A.; Johora, F.T.; Taniya, M.A.; Araf, Y. Immunoinformatics-guided designing of epitope-based subunit vaccines against the SARS Coronavirus-2 (SARS-CoV-2). *Immunobiology* **2020**, *225*, 151955. [[CrossRef](#)]
73. 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](#)]
74. Rehman, I.; Kerndt, C.C.; Botelho, S. *Biochemistry, Tertiary Protein Structure*; StatPearls: Treasure Island, FL, USA, 2021.
75. Enany, S. Structural and functional analysis of hypothetical and conserved proteins of *Clostridium tetani*. *J. Infect. Public Health* **2014**, *7*, 296–307. [[CrossRef](#)]

76. Kyte, J.; Doolittle, R.F. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **1982**, *157*, 105–132. [[CrossRef](#)]
77. Foroutan, M.; Ghaffarifar, F.; Sharifi, Z.; Dalimi, A. Vaccination with a novel multi-epitope ROP8 DNA vaccine against acute Toxoplasma gondii infection induces strong B and T cell responses in mice. *Comp. Immunol. Microbiol. Infect. Dis.* **2020**, *69*, 101413. [[CrossRef](#)]
78. Wlodawer, A. Stereochemistry and Validation of Macromolecular Structures. *Methods Mol. Biol.* **2017**, *1607*, 595–610. [[CrossRef](#)]
79. Yazdani, Z.; Rafiei, A.; Yazdani, M.; Valadan, R. Design an Efficient Multi-Epitope Peptide Vaccine Candidate Against SARS-CoV-2: An in silico Analysis. *Infect. Drug Resist.* **2020**, *13*, 3007–3022. [[CrossRef](#)]
80. Tso, F.Y.; Lidenge, S.J.; Pena, P.B.; Clegg, A.A.; Ngowi, J.R.; Mwaiselage, J.; Ngalamika, O.; Julius, P.; West, J.T.; Wood, C. High prevalence of pre-existing serological cross-reactivity against severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) in sub-Saharan Africa. *Int. J. Infect. Dis.* **2021**, *102*, 577–583. [[CrossRef](#)]
81. Saletti, G.; Gerlach, T.; Jansen, J.M.; Molle, A.; Elbahesh, H.; Ludlow, M.; Li, W.; Bosch, B.J.; Osterhaus, A.; Rimmelzwaan, G.F. Older adults lack SARS CoV-2 cross-reactive T lymphocytes directed to human coronaviruses OC43 and NL63. *Sci. Rep.* **2020**, *10*, 21447. [[CrossRef](#)]
82. Ladner, J.T.; Henson, S.N.; Boyle, A.S.; Engelbrektson, A.L.; Fink, Z.W.; Rahee, F.; D’Ambrozio, J.; Schaecher, K.E.; Stone, M.; Dong, W.; et al. Epitope-resolved profiling of the SARS-CoV-2 antibody response identifies cross-reactivity with an endemic human CoV. *bioRxiv* **2020**. [[CrossRef](#)]
83. Khan, S.; Nakajima, R.; Jain, A.; de Assis, R.R.; Jasinskas, A.; Obiero, J.M.; Adenaiye, O.; Tai, S.; Hong, F.; Milton, D.K.; et al. Analysis of Serologic Cross-Reactivity Between Common Human Coronaviruses and SARS-CoV-2 Using Coronavirus Antigen Microarray. *bioRxiv* **2020**. [[CrossRef](#)]
84. Che, X.Y.; Qiu, L.W.; Liao, Z.Y.; Wang, Y.D.; Wen, K.; Pan, Y.X.; Hao, W.; Mei, Y.B.; Cheng, V.C.; Yuen, K.Y. Antigenic cross-reactivity between severe acute respiratory syndrome-associated coronavirus and human coronaviruses 229E and OC43. *J. Infect. Dis.* **2005**, *191*, 2033–2037. [[CrossRef](#)]
85. Ilinskaya, A.N.; Dobrovolskaia, M.A. Understanding the immunogenicity and antigenicity of nanomaterials: Past, present and future. *Toxicol. Appl. Pharmacol.* **2016**, *299*, 70–77. [[CrossRef](#)]
86. Li, W.; Joshi, M.D.; Singhanian, S.; Ramsey, K.H.; Murthy, A.K. Peptide Vaccine: Progress and Challenges. *Vaccines* **2014**, *2*, 515–536. [[CrossRef](#)]
87. Zhang, H.H.; Yang, X.M.; Xie, Q.M.; Ma, J.Y.; Luo, Y.N.; Cao, Y.C.; Chen, F.; Bi, Y.Z. The potent adjuvant effects of chicken beta-defensin-1 when genetically fused with infectious bursal disease virus VP2 gene. *Vet. Immunol. Immunopathol.* **2010**, *136*, 92–97. [[CrossRef](#)]
88. Lam, L.K.M.; Watson, A.M.; Ryman, K.D.; Klimstra, W.B. Gamma-interferon exerts a critical early restriction on replication and dissemination of yellow fever virus vaccine strain 17D-204. *NPJ Vaccines* **2018**, *3*, 5. [[CrossRef](#)]
89. Ewer, K.J.; Barrett, J.R.; Belij-Rammerstorfer, S.; Sharpe, H.; Makinson, R.; Morter, R.; Flaxman, A.; Wright, D.; Bellamy, D.; Bittaye, M.; et al. T cell and antibody responses induced by a single dose of ChAdOx1 nCoV-19 (AZD1222) vaccine in a phase 1/2 clinical trial. *Nat. Med.* **2021**, *27*, 270–278. [[CrossRef](#)]
90. Gallagher, K.M.E.; Leick, M.B.; Larson, R.C.; Berger, T.R.; Katsis, K.; Yam, J.Y.; Brini, G.; Grauwet, K.; MGH COVID-19 Collection & Processing Team; Maus, M.V. SARS-CoV-2 T-cell immunity to variants of concern following vaccination. *bioRxiv* **2021**. [[CrossRef](#)]
91. Dong, R.; Chu, Z.; Yu, F.; Zha, Y. Contriving Multi-Epitope Subunit of Vaccine for COVID-19: Immunoinformatics Approaches. *Front. Immunol.* **2020**, *11*, 1784. [[CrossRef](#)]
92. Zhao, Y.; Kuang, M.; Li, J.; Zhu, L.; Jia, Z.; Guo, X.; Hu, Y.; Kong, J.; Yin, H.; Wang, X.; et al. SARS-CoV-2 spike protein interacts with and activates TLR41. *Cell Res.* **2021**. [[CrossRef](#)]
93. Tahir Ul Qamar, M.; Rehman, A.; Tusleem, K.; Ashfaq, U.A.; Qasim, M.; Zhu, X.; Fatima, I.; Shahid, F.; Chen, L.L. Designing of a next generation multiepitope based vaccine (MEV) against SARS-COV-2: Immunoinformatics and in silico approaches. *PLoS ONE* **2020**, *15*, e0244176. [[CrossRef](#)]
94. Crooke, S.N.; Ovsyannikova, I.G.; Kennedy, R.B.; Poland, G.A. Immunoinformatic identification of B cell and T cell epitopes in the SARS-CoV-2 proteome. *Sci. Rep.* **2020**, *10*, 14179. [[CrossRef](#)]
95. 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](#)]
96. Mukherjee, S.; Tworowski, D.; Detroja, R.; Mukherjee, S.B.; Frenkel-Morgenstern, M. Immunoinformatics and Structural Analysis for Identification of Immunodominant Epitopes in SARS-CoV-2 as Potential Vaccine Targets. *Vaccines* **2020**, *8*, 290. [[CrossRef](#)]
97. Wang, D.; Mai, J.; Zhou, W.; Yu, W.; Zhan, Y.; Wang, N.; Epstein, N.D.; Yang, Y. Immunoinformatic Analysis of T- and B-Cell Epitopes for SARS-CoV-2 Vaccine Design. *Vaccines* **2020**, *8*, 355. [[CrossRef](#)]
98. Toledo, H.; Baly, A.; Castro, O.; Resik, S.; Laferte, J.; Rolo, F.; Navea, L.; Lobaina, L.; Cruz, O.; Miguez, J.; et al. A phase I clinical trial of a multi-epitope polypeptide TAB9 combined with Montanide ISA 720 adjuvant in non-HIV-1 infected human volunteers. *Vaccine* **2001**, *19*, 4328–4336. [[CrossRef](#)]
99. Slingluff, C.L., Jr.; Lee, S.; Zhao, F.; Chianese-Bullock, K.A.; Olson, W.C.; Butterfield, L.H.; Whiteside, T.L.; Leming, P.D.; Kirkwood, J.M. A randomized phase II trial of multiepitope vaccination with melanoma peptides for cytotoxic T cells and helper T cells for patients with metastatic melanoma (E1602). *Clin. Cancer Res.* **2013**, *19*, 4228–4238. [[CrossRef](#)]

100. Lennerz, V.; Gross, S.; Gallerani, E.; Sessa, C.; Mach, N.; Boehm, S.; Hess, D.; von Boehmer, L.; Knuth, A.; Ochsenbein, A.F.; et al. Immunologic response to the survivin-derived multi-epitope vaccine EMD640744 in patients with advanced solid tumors. *Cancer Immunol. Immunother.* **2014**, *63*, 381–394. [[CrossRef](#)]
101. Guo, L.; Yin, R.; Liu, K.; Lv, X.; Li, Y.; Duan, X.; Chu, Y.; Xi, T.; Xing, Y. Immunological features and efficacy of a multi-epitope vaccine CTB-UE against *H. pylori* in BALB/c mice model. *Appl. Microbiol. Biotechnol.* **2014**, *98*, 3495–3507. [[CrossRef](#)]
102. Cao, Y.; Li, D.; Fu, Y.; Bai, Q.; Chen, Y.; Bai, X.; Jing, Z.; Sun, P.; Bao, H.; Li, P.; et al. Rational design and efficacy of a multi-epitope recombinant protein vaccine against foot-and-mouth disease virus serotype A in pigs. *Antiviral. Res.* **2017**, *140*, 133–141. [[CrossRef](#)]