



*biomedicines*



Article

---

# An Immunoinformatic Approach for Identifying and Designing Conserved Multi-Epitope Vaccines for Coronaviruses

---

Yu Chuan Ong, Bimo Ario Tejo and Wei Boon Yap

Topic

Advances in Vaccines and Antimicrobial Therapy

Edited by

Dr. Roberto Paganelli and Dr. Raffaele D'Amelio



<https://doi.org/10.3390/biomedicines12112530>



## Article

# An Immunoinformatic Approach for Identifying and Designing Conserved Multi-Epitope Vaccines for Coronaviruses

Yu Chuan Ong <sup>1</sup>, Bimo Ario Tejo <sup>2</sup> and Wei Boon Yap <sup>1,3,\*</sup>

<sup>1</sup> Center for Toxicology and Health Risk Studies, Faculty of Health Sciences, Universiti Kebangsaan Malaysia, Kuala Lumpur 50300, Malaysia; p131331@siswa.ukm.edu.my

<sup>2</sup> Department of Chemistry, Faculty of Science, Universiti Putra Malaysia, Serdang 43400, Malaysia; bimo.tejo@upm.edu.my

<sup>3</sup> One Health UKM, Universiti Kebangsaan Malaysia, Bangi 43600, Malaysia

\* Correspondence: yapweiboon@ukm.edu.my; Tel.: +603-92897920; Fax: +603-26929032

**Abstract:** Background/Objectives: The COVID-19 pandemic caused by the novel severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) virus has exposed the vulnerabilities and unpreparedness of the global healthcare system in dealing with emerging zoonoses. In the past two decades, coronaviruses (CoV) have been responsible for three major viral outbreaks, and the likelihood of future outbreaks caused by these viruses is high and nearly inevitable. Therefore, effective prophylactic universal vaccines targeting multiple circulating and emerging coronavirus strains are warranted. Methods: This study utilized an immunoinformatic approach to identify evolutionarily conserved CD4+ (HTL) and CD8+ (CTL) T cells, and B-cell epitopes in the coronaviral spike (S) glycoprotein. Results: A total of 132 epitopes were identified, with the majority of them found to be conserved across the bat CoVs, pangolin CoVs, endemic coronaviruses, SARS-CoV-2, and Middle East respiratory syndrome coronavirus (MERS-CoV). Their peptide sequences were then aligned and assembled to identify the overlapping regions. Eventually, two major peptide assemblies were derived based on their promising immune-stimulating properties. Conclusions: In this light, they can serve as lead candidates for universal coronavirus vaccine development, particularly in the search for pan-coronavirus multi-epitope universal vaccines that can confer protection against current and novel coronaviruses.



**Citation:** Ong, Y.C.; Tejo, B.A.; Yap, W.B. An Immunoinformatic Approach for Identifying and Designing Conserved Multi-Epitope Vaccines for Coronaviruses. *Biomedicines* **2024**, *12*, 2530. <https://doi.org/10.3390/biomedicines12112530>

Academic Editors: Roberto Paganelli and Raffaele D'Amelio

Received: 6 September 2024

Revised: 30 October 2024

Accepted: 2 November 2024

Published: 5 November 2024



**Copyright:** © 2024 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/>).

**Keywords:** SARS-CoV-2; coronavirus; universal vaccine; peptide; cell-mediated immunity

## 1. Introduction

Coronaviruses (CoVs) were thought to only cause common colds in humans until 2002–2003, when severe acute respiratory syndrome-coronavirus (SARS-CoV) struck the world health system. The SARS-CoV outbreak caused 8098 infections and 774 deaths globally (~10% mortality) [1]. Nearly 10 years after the SARS outbreak, another coronavirus outbreak took place in the Middle East and the etiological agent was identified as Middle East respiratory syndrome coronavirus (MERS-CoV). The outbreak then spread to South Korea, with the first case reported in 2015 involving a 65-year-old man who had recently travelled to the Middle East [2]. Compared to SARS-CoV, MERS-CoV had the highest mortality rate of ~35% among the three outbreak strains, with a total of 2458 infections and 848 reported deaths. In December 2019, a novel coronavirus strain namely SARS-CoV-2 caused a large-scale viral outbreak in WuHan, China. Since then, it has spread rampantly throughout the world and caused respiratory distress in humans. The disease associated with SARS-CoV-2 was coined COVID-19 and resulted in the breakdown of the healthcare system in many countries. The outbreak was eventually announced as a global pandemic by WHO on 11 March 2020 [3]. As of July 2024, COVID-19 has affected more than 775 million people and caused more than 7 million deaths worldwide [4].

As the coronaviral spike (S) glycoprotein is located outside the viral particle and mediates viral entry into host epithelial cells, it is undoubtedly the main target of neutralizing

antibodies (NAbs) upon infection, making it the most important therapeutic target and the main focus in vaccine design. However, the emergence of new SARS-CoV-2 Omicron variants has rendered the vaccines ineffective, with ChAdOx1 nCoV-19 (Vaxzevria, AstraZeneca) conferring almost no protection from 20–24 weeks after the second dose of vaccine [5]. It is also notable that the emergence of new variants of concern (VOC), such as Omicron, has attracted attention globally as the new variants can escape the neutralizing antibodies and have increased transmissibility due to the presence of more than 30 mutations as compared to the parental strain, SARS-CoV-2-Wuhan-Hu-1 [6–9]. In view of the rising concerns regarding the increased infectivity of the new variants and controversies about the effectiveness of the existing vaccines, there is an urgent need for a pan-coronavirus vaccine that can induce the synthesis of neutralizing antibodies and is more comprehensive in conferring protection against the newly emerging variants as well as future coronavirus outbreaks. While many groups have predicted and identified evolutionarily conserved epitopes in silico [10–16], and some of them were validated in vitro and in vivo [17–23], this study scrutinized the conserved epitopes further. Many predicted cytotoxic T lymphocyte (CTL), helper T lymphocyte (HTL), and linear B lymphocyte (LBL) epitopes were found in close vicinity to the S glycoprotein. Instead of studying them individually, they were aligned into single and relatively long peptide sequences. This novel strategy of having multi-epitopes is expected to stimulate a stronger and multi-faceted immune response against coronaviruses, addressing the limitations of the current vaccines against the emerging variants.

In this study, the evolutionarily conserved epitopes in both human and animal coronaviruses were identified using unique immunoinformatic approaches. After a stringent scrutiny and selection, 52 CTL epitopes, 11 HTL epitopes, and 68 linear B-lymphocyte (LBL) epitopes were identified from 30 coronavirus sequences derived from human CoVs (hCoVs) responsible for the common cold, SARS-CoV, MERS-CoV and SARS-CoV-2. Subsequently, the predicted epitopes were aligned and assembled into two final composite peptide sequences that were found to be evolutionarily conserved across SARS-CoVs, bats, and pangolin coronaviruses. These two assembled epitopes were not only found to be conserved in many of the coronavirus strains, but they also possessed HTL, CTL, and B-cell antigen binding sites, and they matched a diverse array of HLA class I and II supertypes prevalent in the human population, indicating their potential to activate both T and B cells effectively. Although these epitopes were identified on the basis of being conserved in SARS-CoVs and bat and pangolin coronaviruses, their distinctive compatibility with human HTL and CTL and B cells renders them of high potential in vaccine development. Altogether, these discoveries not only pave the way for the development of a pan-coronavirus multi-epitope vaccine to combat existing and novel coronavirus strains but its immunoinformatics are highly applicable in universal vaccine development, especially in identifying immunogenic conserved epitopes in target antigens.

## 2. Materials and Methods

### 2.1. Coronaviral S Gene Sequence Retrieval and Sequence Conservation Analysis

Forty-two coronaviral peptide sequences of the S gene were retrieved from the NCBI GenBank (<https://www.ncbi.nlm.nih.gov/genbank/>) (accessed on 28 July 2024) as listed in Tables 1–3. A total of 24 sequences encompassing SARS-CoV-2 and its variants were retrieved from the NCBI GenBank (Table 1) with the latest variant being XBB.1.5. Six sequences of the other coronaviruses causing SARS, MERS and common colds in humans are listed in Table 2, and Table 3, on the other hand, tabulates twelve sequences of coronaviruses isolated from bats, pangolins and birds.

**Table 1.** SARS-CoV-2 and its variants.

Nomenclature	Lineage	Accession Number
SARS-CoV-2-Wuhan-Hu-1 strain		NC_045512.2
Alpha	B.1.1.7	OK340744.1
Beta	B.1.351	OQ341818.1
Delta	B.1.617.2	OQ314763.1
Gamma	P.1	OQ316323.1
Omicron	B.1.1.529	OQ344199.1
Omicron	BA.1	OQ355083.1
Omicron	BA.1.1	OQ352636.1
Omicron	BA.2	OQ341824.1
Omicron	BA.2.12.1	OQ355080.1
Omicron	BA.2.75	OQ215893.1
Omicron	BA.2.75.2	OQ346937.1
Omicron	BA.4	OQ333888.1
Omicron	BA.4.6	OQ349323.1
Omicron	BA.5	OQ343976.1
Omicron	BA.5.2.6	OQ346806.1
Omicron	BF.11	OQ347094.1
Omicron	BF.7	OQ346784.1
Omicron	BN.1	OQ346744.1
Omicron	BQ.1	OQ346454.1
Omicron	BQ.1.1	OQ346605.1
Omicron	CH.1.1	OQ346876.1
Omicron	XBB	OQ347865.1
Omicron	XBB.1.5	XBB.1.5 is a sub-lineage of XBB with an additional spike RBD mutation S486P

**Table 2.** Sequences of hCoVs, SARS and MERS with their accession numbers.

Nomenclature	Accession Number
MERS-CoV	NC_019843
SARS-CoV (Urbani)	AY278741.1
HCoV-HKU1–genotype B	AY884001
HCoV-OC43	KF923903
HCoV-NL63	NC_005831

**Table 3.** Coronaviruses infecting bats, pangolins and birds.

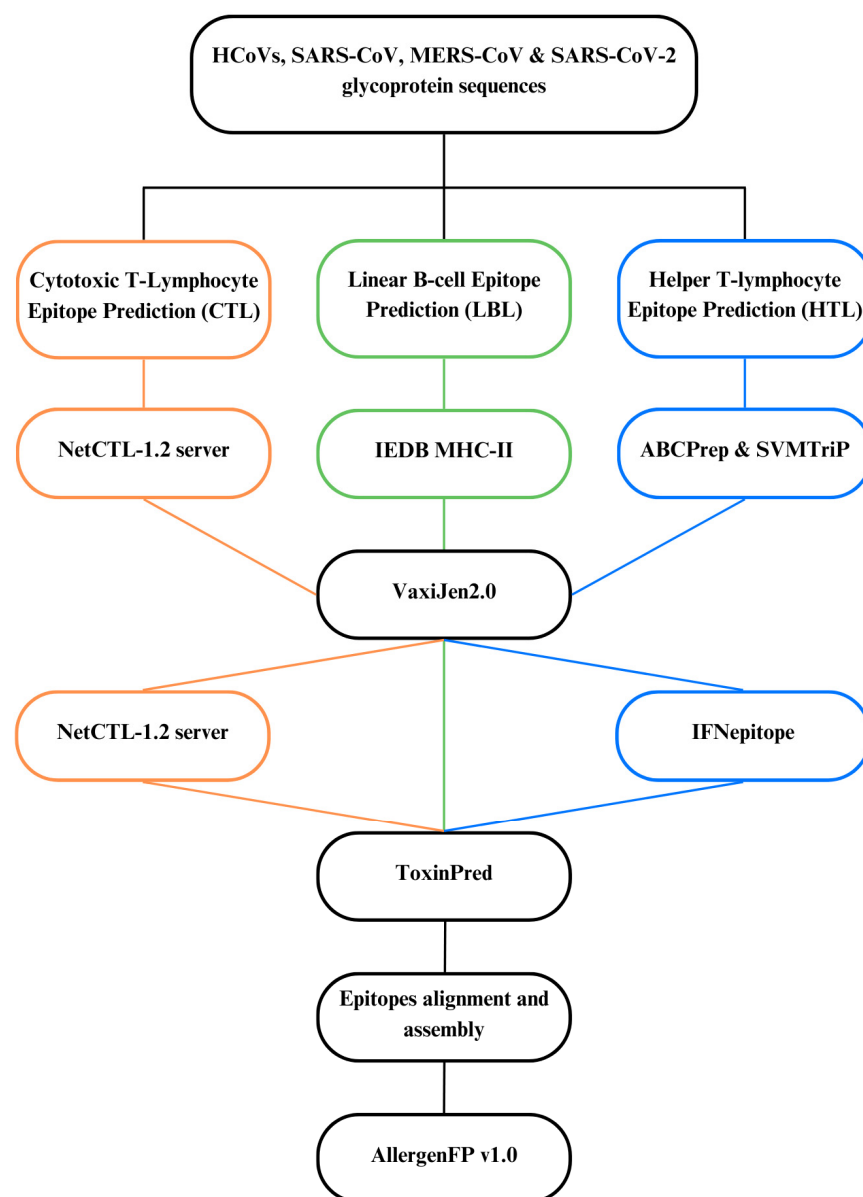
Strain Name	Accession Number
Bat CoV RATG13	MN996532.2
Bat CoV ZXC21	MG772934.1
Bat CoV YN02	MW201982.1
Pangolin CoV GX-P2V	MT072864.1
Pangolin CoV GX-P5E	MT040336.1
Pangolin CoV GX-P5L	MT040335.1
Pangolin CoV GX-P1E	MT040334.1
Pangolin CoV GX-P4L	MT040333.1
Pangolin CoV MP789	MT121216.1
Avian CoV Ind-TN92-03	NC_048213.1
Avian CoV DK/GD/27/2014	NC_048214.1
Avian CoV MG10	NC_010800.1

In order to identify the conserved regions in the coronaviral S glycoproteins, the amino acid sequence of the S glycoprotein of SARS-CoV-2 Wuhan-Hu-1 strain (Table 1) was used as a reference sequence to perform Clustal Omega multiple sequence alignments

in the EMBL-EBI (<https://www.ebi.ac.uk/> (accessed on 28 July 2024)). The alignment was based on the Percentage Identity Threshold of 80% in the amino acid sequences using Jalview 2.11.2.6 (<https://www.jalview.org/> (accessed on 28 July 2024)). The evolutionarily conserved regions of the S glycoproteins were identified and subjected to antigenicity screening, selection, and assembly.

## 2.2. The Flow of Prediction of Conserved HTL, CTL and Linear B-Lymphocyte (LBL) Epitopes of Coronaviral S Glycoproteins

The prediction was performed separately for (i) CTL, (ii) HTL, and (iii) LBL epitopes by referring to their respective databases. The flow of prediction of the conserved epitopes is depicted in Figure 1. The conserved epitopes were individually screened and identified, and their antigenicity and toxicity were predicted using VaxiJen 2.0 and ToxinPred, respectively.



**Figure 1.** Flow of in silico prediction of conserved epitopes of the coronaviral S proteins. The orange, green and blue-colored lines represent the CTL, LBL and HTL prediction steps, respectively.

### 2.2.1. Prediction of Conserved CTL Epitopes

The conserved CTL epitopes were identified using NetCTL-1.2 (<https://services.healthtech.dtu.dk/services/NetCTL-1.2/>) (accessed on 1 August 2024). A total of 30 amino acid sequences of human coronaviral S glycoproteins were uploaded to NetCTL-1.2 by following the default criteria, which entailed 9 amino acids in length with a minimum threshold of 0.75. The available HLA class I supertypes provided by NetCTL-1.2 included A1, A2, A24, A26, B7, B8, B27, B39, B44, B58, and B62. The redundant epitope sequences were filtered and subjected to the subsequent screening. The selected epitopes were then subjected to in silico antigenicity screening using VaxiJen 2.0 (<http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html>) (accessed on 1 August 2024). The screening was performed using the default settings and “Virus” was selected as the target organism as part of the prediction criteria. Antigens labeled as “Probable Antigen” were sorted and selected. Subsequently, the selected antigens were screened for their immunogenicity using the IEDB Class I immunogenicity web-based prediction tool (<http://tools.iedb.org/immunogenicity/>) (accessed on 1 August 2024). The non-immunogenic epitopes were indicated with negative scores and removed and the remaining epitopes were later examined using ToxinPred (<https://webs.iitd.edu.in/raghava/toxinpred/protein.php>) (accessed on 1 August 2024) to eliminate the probable toxic CTL epitopes. Following that, the epitopes were selected based on their high frequencies across different strains and supertypes. Table 4 summarizes the prediction of CTL epitopes.

**Table 4.** CTL prediction tools and their prediction criteria.

CTL Prediction Tools	Prediction Tool's Criteria
NetCTL-1.2	<ol style="list-style-type: none"> <li>1. Threshold: 0.75, 9-mers.</li> <li>2. Predict with all available supertypes (A1, A2, A24, A26, B7, B8, B27, B39, B44, B58, and B62).</li> <li>3. Select sequences with a combined score of above 0.75.</li> <li>4. Remove repetitive epitope(s) after prediction.</li> </ol>
VaxiJen 2.0	<ol style="list-style-type: none"> <li>1. Target Organism: Virus.</li> <li>2. Threshold: Default.</li> <li>3. Exclude “Non-antigenic” epitope(s).</li> </ol>
IEDB MHC Class I immunogenicity	<ol style="list-style-type: none"> <li>1. Masking position: Default.</li> <li>2. Exclude non-immunogenic epitopes.</li> <li>3. Prediction method: SVM (Swiss-Prot) based.</li> </ol>
ToxinPred	<ol style="list-style-type: none"> <li>1. Quantitative Matrix (QM) method: Blank.</li> <li>2. E-value cut-off for motif-based method: 10.</li> <li>3. SVM threshold: 0.</li> <li>4. Exclude “Toxin” epitope(s)</li> </ol>

### 2.2.2. Prediction of Conserved HTL Epitopes

In the conserved HTL epitope prediction, the amino acid sequences of the coronaviruses were screened using IEDB MHC-II (<http://tools.iedb.org/mhcii/>) (accessed on 1 August 2024) with the following prediction conditions: (i) Percentile rank: 25%, 15-mers amino acid; (ii) prediction method: Consensus 2.22; (iii) HLA supertypes: HLA-DR, HLA-DQ, and HLA-DP. Table 5 summarizes the prediction conditions for the conserved HTL epitopes. Epitopes with a percentile rank lower than 20.0 were eliminated as this indicates that those epitopes capture a less than 50% immune response [24].

Similar to the conserved CTL epitope prediction, the antigenicity and toxigenicity of the epitopes were analyzed using the same methods as shown in Table 4. In addition, the HTL epitopes were screened using IFNepitope (<https://webs.iitd.edu.in/raghava/ifnepitope/predict.php>) (accessed on 1 August 2024) for their abilities to induce interferon synthesis. The prediction criteria included “Motif and SVM hybrid” as the prediction

approach and “IFN-gamma versus Non IFN-gamma” as the prediction model. The “NEGATIVE” HTL epitopes were removed.

**Table 5.** Prediction flow and criteria of conserved HTL epitopes.

HTL Prediction Tools	Prediction Tool’s Criteria
IEDB MHC-II	1. Percentile rank: 20%, 15-mers.
	2. Method: Consensus 2.22.
	3. HLA Supertype: HLA-DR, HLA-DQ, HLA-DP.
	i. HLA-DR: <ul style="list-style-type: none"> <li>• DRB1*01:01</li> <li>• DRB1*07:01</li> <li>• DRB1*09:01</li> <li>• DRB3-01:01</li> <li>• DRB4*01:01</li> </ul>
IFNepitope	ii. HLA-DQ: <ul style="list-style-type: none"> <li>• DQA1*01:01/DQB1*05:01</li> <li>• DQA1*01:02/DQB1*06:02</li> <li>• DQA1*03:01/DQB1*03:02</li> <li>• DQA1*04:01/DQB1*04:02</li> <li>• DQA1*05:01/DQB1*02:01</li> <li>• DQA1*05:01/DQB1*03:01</li> </ul>
	iii. HLA-DP: <ul style="list-style-type: none"> <li>• DPA1*01/DPB1*04:01</li> <li>• DPA1*01:03/DPB1*02:01</li> <li>• DPA1*02:01/DPB1*01:01</li> <li>• DPA1*02:01/DPB1*05:01</li> <li>• DPA1*03:01/DPB1*04:02</li> </ul>
	4. Exclude epitope(s) with percentile rank higher than 20.0
	1. Prediction approach: Motif and SVM hybrid.
	2. Model for prediction: IFN-gamma versus non IFN-gamma.
	3. Exclude “NEGATIVE” epitope(s).

2.2.3. Prediction of Conserved LBL Epitopes

ABCPred ([https://webs.iitd.edu.in/raghava/abcpred/ABC\\_submission.html](https://webs.iitd.edu.in/raghava/abcpred/ABC_submission.html) (accessed on 1 August 2024)) and SVMTriP (<http://sysbio.unl.edu/SVMTriP/prediction.php> (accessed on 1 August 2024)) were used for predicting the conserved LBL epitopes with criteria such as 16-mers amino acids in length for both tools and the thresholds of 0.51 and 0.50, respectively, as shown in Table 6. All of the predicted epitopes were selected for further antigenicity and toxigenicity prediction. The antigenicity and toxigenicity of the epitopes were screened as described in Table 4.

**Table 6.** Prediction flow and criteria of conserved LBL epitopes.

LBL Prediction Tools	Prediction Tool’s Condition
ABCPred	1. Length of epitope: 16-mers
	2. Threshold: 0.51 and above
	3. Overlapping filter: ON
SVMTriP	1. Length of epitope: 16-mers
	2. Select epitope(s) with a score of 0.5 and above

2.3. Alignment of the Predicted Conserved CTL, HTL and LBL Epitopes and Allergenicity Prediction

To identify the locations of the epitopes identified in Section 2.2, SARS-CoV-2-Wuhan-Hu-1 S glycoprotein was used as a reference sequence for multiple sequence alignment.



The overlapped regions of the epitopes were aligned. Then, they were assembled into long amino acid sequences containing the conserved CTL, HTL, and LBL epitopes. Subsequently, the assembled peptide sequences were screened for allergenicity using AllergenFP v1.0 (<https://www.ddg-pharmfac.net/AllergenFP/> (accessed on 1 August 2024)) to identify the probable allergens in the assembled sequences. The probable allergenic sequences, if any, were eliminated.

#### 2.4. Population Coverage Analysis

The assembled epitopes were subjected to population coverage analysis using the IEDB Population Coverage analysis tool (<http://tools.iedb.org/population/> (accessed on 5 October 2024)) with the default parameters. The analysis was conducted based on the following parameters: (i) “World” as the selected area/population, and (ii) the calculation was determined for “Class I”, “Class II”, and “Class I and II combined”. The “MHC-restricted allele(s)” were determined based on the predicted HLA Class I and Class II associated with the assembled epitopes.

#### 2.5. Structural Visualisation of Assembled Epitopes

The structural information of the closed (PDB: 6VXX) and open states (PDB: 6VYB) of the S glycoprotein were retrieved from the RCSB Protein Data Bank (<https://www.rcsb.org/> (accessed on 10 August 2024)). By using the retrieved information, the structures of the assembled conserved epitopes were then visualized using UCSF ChimeraX (<https://www.cgl.ucsf.edu/chimerax/> (accessed on 10 August 2024)).

#### 2.6. Molecular Docking of the Assembled Epitopes to TLR4 and TLR2 Receptors

The molecular docking was performed using HADDOCK 2.4 (<https://www.bonvinlab.org/software/haddock2.4/> (accessed on 15 October 2024)) to determine the interaction of the assembled epitopes with cell receptors involved in the vaccine-induced immune responses. Specifically, both Toll-Like Receptor 2 (TLR2) (PDB ID: 2z7x) and TLR4 dimer (PDB ID: 4g8a) were selected as target receptors due to their localization on the cell surface and their roles in initiating innate immune responses upon binding to vaccine components [25,26]. The binding was analyzed using PDBsum (<https://www.ebi.ac.uk/thornton-srv/databases/pdbsum/> (accessed on 19 October 2024)) and the binding affinity in terms of Gibbs free energy ( $\Delta G$ ) and dissociation constants (M) at 25 °C between the docked molecules were calculated using the PRODIGY web server (<https://rascar.science.uu.nl/prodigy/> (accessed on 19 October 2024)) [27,28].

#### 2.7. Immune Simulation Using C-IMMSIM Server

Immune simulation was performed on the C-IMMSIM server (<https://kraken.iac.rm.cnr.it/C-IMMSIM/> (accessed on 19 October 2024)) to characterize the immune response profile and immunogenicity of the assembled peptides [29]. The entire simulation was performed for 300-time steps, equivalent to about 80 days (a time step is about 8 h). Three peptide injections were given four weeks apart at time steps 10, 94, 178 [29].

### 3. Results

#### 3.1. Conserved Regions in the S Glycoproteins of Bat and Pangolin CoV, hCoV, SARS-CoV-2, SARS-CoV, and MERS-CoV

The S glycoprotein of coronaviruses consists of a signal peptide, a receptor binding subunit (S1) and a fusion subunit (S2) with a length ranging from 1105–1351 amino acids [30–33]. With reference to the SARS-CoV-2-Wuhan-Hu-1, in the S1 subunit, there is an N-terminal domain (NTD, 14–305 residues) and receptor-binding domain (RBD, 319–541 residues), whereas, in the S2 subunit, there is a fusion peptide (FP, 788–806 residues), heptapeptide repeat sequence 1 (HR1) (912–984 residues), HR2 (1163–1213 residues), transmembrane domain (TM, 1213–1237 residues), and cytoplasmic domain (1237–1273 residues) [34]. As the coronavirus S glycoprotein is located outside of the viral particle and mediates the



viral entry into the host epithelial cells, it is undoubtedly the main target of neutralizing antibodies (NAbs) upon infection, making it the most important therapeutic target and essential in vaccine design.

Given the importance of coronaviral S glycoprotein in vaccine development, this study aimed to screen and identify the evolutionarily conserved sequences in animal and human coronaviral S glycoproteins. The amino acid sequence of SARS-CoV-2-Wuhan-Hu-1 S glycoprotein served as the reference sequence for all retrieved coronavirus sequences. Upon screening and alignment, the results revealed that bat coronavirus strains, such as Bat CoV RATG13, Bat CoV ZXC21, and Bat CoV YN02, and pangolin coronavirus strains, such as Pangolin CoV GX-P2V, Pangolin CoV GX-P5E, Pangolin CoV GX-P5L, Pangolin CoV GX-P1E, Pangolin CoV GX-P4L, and Pangolin CoV MP789, showed some level of evolutionary divergence compared to that of SARS-CoV-2-Wuhan-Hu-1 with an identity threshold above 80% (Supplementary Figure S1). These results also explain why bats or pangolins are deduced as the most likely reservoirs of SARS-CoV-2. In contrast, three selected avian coronaviral S glycoproteins were distantly related to the reference sequence due to their relatively high evolutionary divergence.

The amino acid sequences of SARS-CoV-2 and hCoVs, causing the common cold, were also aligned with the reference sequence (Supplementary Figure S2). The S1 regions of hCoVs, i.e., H-CoV-HKU1-genotype B, CoV-OC43, CoV-NL63, and CoV-229E, showed insignificant similarities to the reference sequence. Interestingly, their S2 regions were relatively conserved, particularly at residues S815–S874, S897–S934, S944–S1069, and S1207–S1218. The residues S897–S1069 corresponded to the HR1 and HR2 regions of the S2 subunit, whereas the S1207–S1218 region was part of the HR2 and TM domain. The conservation of the HR1 and HR2 regions was documented previously and suggested as the targets for the development of fusion inhibitor agents [35,36]. Furthermore, the alignment of S glycoprotein sequences of SARS-CoV-2 and its variants, MERS-CoV and SARS-CoV, revealed that SARS-CoV, SARS-CoV-2 and its variants were highly similar to each other (Supplementary Figure S3). This finding suggests that the emergence of SARS-CoV and SARS-CoV-2 might be due to the recombination of viral genomes between bat coronaviruses in their natural reservoir (bats) or the intermediate host (pangolin), or both. There were no observable evolutionarily conserved regions in the S glycoprotein sequence of MERS-CoV relative to that of SARS-CoV. Altogether, the alignment of coronaviral S glycoproteins with the reference sequence revealed a high evolutionary relationship between SARS-CoV (Urbani), bat CoVs, and pangolin CoVs. It is suggested that the emergence of the highly contagious and pandemic-causing SARS-CoV-2 is highly attributable to genome recombination or mutations of the coronavirus in animal hosts such as bats [37–39]. The high evolutionary relationship among coronaviruses sheds light on the development of universal vaccines using conserved epitopes.

### 3.2. Prediction and Screening of Conserved CTL Epitopes of S Glycoprotein

The conserved CTL epitopes of S glycoprotein were first screened and predicted based on 30 coronaviral S glycoprotein sequences. The NetCTL-1.2 of DTU Health Tech provides high sensitivity and specificity among the publicly available bioinformatics tools [40,41]. This web-based bioinformatics tool utilizes a combination of predictive algorithms including proteasomal cleavage, TAP transport efficiency, and MHC class I affinity to acquire highly probable CTL epitopes in a given sequence. Given the easy accessibility, the epitopes were selected based on the available human leucocyte antigen (HLA) class I supertypes provided by the algorithms, such as A1, A2, A3, A24, A26, B7, B8, B27, B39, B44, B58, and B62 supertypes.

HLA class I is known to be responsible for presenting processed antigens to T-cell receptors. Generally, there are three classical HLA class I encoding genes (HLA-A, HLA-B, and HLA-C) and all of them are extremely polymorphic. The number of identified HLA alleles has grown exponentially over the past decades and is likely to increase with time. To date, there are over 36,000 sequences of highly curated HLA alleles deposited in the

IPD-IMGT/HLA Database (<https://www.ebi.ac.uk/ipd/imgt/hla/> (accessed on 14 August 2024)). Undoubtedly, the vast number of HLA alleles makes the epitope prediction significantly complex and impractical. Thus, in the mid-1990s, an allele-specific classification called HLA supertype was created, in which the first nine HLA class I superotypes were described [42] and three more HLA class I superotypes were added later. Hence, there are 12 HLA class I superotypes in the latest update [43]. dos Santos Francisco et al. (2015) investigated HLA class I supertype frequencies among 55 human populations and found that HLA superotypes A2, A3, B7, B27, and B44 were evenly distributed and not specific to only certain populations [37]. Half of the populations showed frequencies at 14–29% for A2, 14–32% for A3, 18–31% for B7, and 21–32% for B44. In contrast, HLA superotypes A1, A24, B58, and B62 had greater frequency variations among the studied populations. It is also worth mentioning that the A24 supertype was found at higher frequencies (40% on average) in SEA, PAC, AUS, NEA and AME; meanwhile, the A1 supertype had an average frequency of 21% in Africa, Europe, and Southwest Asia [44].

The prediction of conserved CTL epitopes was based on the HLA class I superotypes to cover as many human populations as possible. The initial screening yielded 1,048,575 potential CTL epitopes that matched the 12 HLA class I superotypes. The large number of epitopes was then streamed down based on their antigenicity, immunogenicity, and toxigenicity. The elimination was performed using VaxiJen 2.0, IEDB MHC Class I immunogenicity and ToxinPred, respectively. The remaining 2114 epitopes were subjected to another round of screening based on their frequency of appearance in 30 coronavirus strains and 12 HLA class I superotypes. After stringent selection and removal of redundant epitopes, 12 epitopes (Table 7) were chosen for further analysis during the epitope's alignment step.

**Table 7.** Final selected CTL epitopes.

Epitopes	Number of Coronavirus Strains in Which the Epitope Is Found (Out of 30)	Location in the S Glycoprotein *	Assigned Name
RVVVLSFEL	25	509–517	CTL1
STQDLFLPF	24	50–59	CTL2
WTAGAAAYY	24	258–266	CTL3
YLPRTFLL	24	269–277	CTL4
QITTDNTF	24	1113–1121	CTL5
GAAAYVGY	24	261–269	CTL6
ITDAVDCAL	24	284–293	CTL7
FTISVTTEI	24	718–726	CTL8
FVFLVLLPL	23	2–9	CTL9
QSYGFRPTY	15	493–501	CTL10
SVLYNFAPF	13	366–374	CTL11
YQPYRVVVL	6	505–513	CTL12

\* Reference sequence: SARS-CoV-2-Wuhan-Hu-1 sequence.

### 3.3. Prediction and Screening of Conserved HTL Epitopes

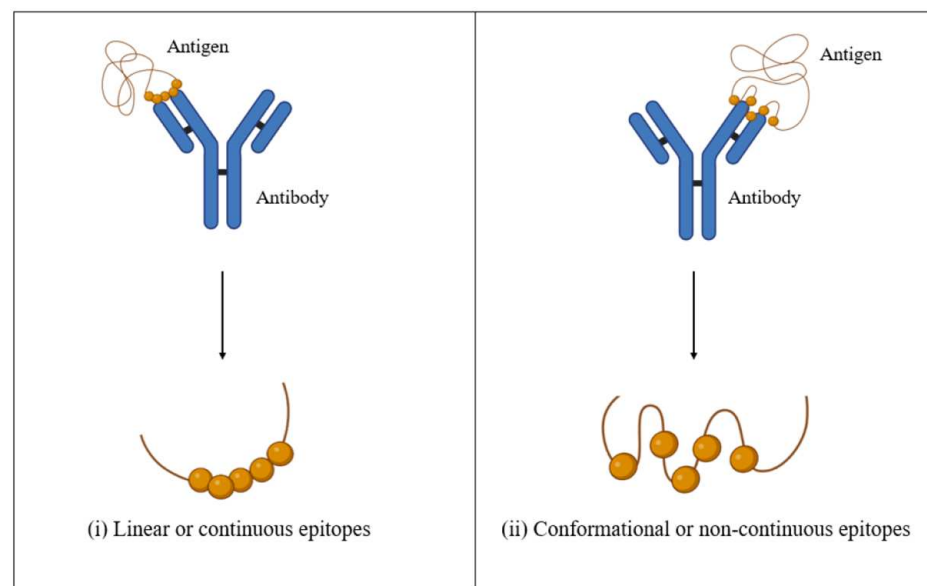
The IEDB MHC-II prediction tool was used to predict and identify HTL epitopes because it provides a remarkable performance score owing to the embedded IEDB consensus 2.22 method [45,46]. Thirty coronavirus S glycoprotein sequences were mapped to 27 most widely distributed HLA class II alleles as described by Greenbaum et al. (2011) [47]. A total of 108,767 HTL epitopes were identified and selected based on their being in the top 20% of the consensus percentile rank, corresponding to their abilities to capture 50% of the total immune response [24].

The epitopes were subjected to antigenicity, IFN-inducing, and toxigenicity predictions. The total number of remaining peptides was 1377, which rendered difficulties in epitope selection. Consequently, peptides with 50% or greater matching with the HLA class II alleles and coronavirus strains were chosen. This is to ensure a wider HLA supertype coverage and provide a more comprehensive defense against multiple strains of coronaviruses. There were 52 epitopes retained (Supplementary Table S1), which were subsequently subjected to

sequence alignment with that of SARS-CoV-2-Wuhan-Hu-1 to locate their positions. All of them were highly related to SARS-CoV-2 and its variants. Among them, three epitopes including HTL3, HTL6 and HTL26 were also related to SARS-CoV (Urbani).

### 3.4. Prediction and Screening of Conserved LBL Epitopes

The identified conserved LBL epitopes represented potential antigen candidates for stimulating the humoral immune response. Generally, B-cell epitopes are divided into (i) linear and continuous or (ii) conformational and non-continuous (Figure 2). Although the vast majority of B-cell epitopes are conformational (approximately 90%) [48,49], the prediction of conformational B-cell epitopes is not as established as the LBL epitopes. Thus, the LBL epitope prediction has gained the most attention, especially in epitope-based vaccine development.



**Figure 2.** Schematic diagram of linear and conformational B-cell epitopes. Panels (i) linear or continuous B-cell epitopes composed of amino acid residues that are sequential to one another; (ii) conformational or non-continuous B-cell epitopes composed of amino acids that are non-sequential and scattered along the peptide sequence.

In contrast to the conserved HTL and CTL epitope predictions, screening and prediction of LBL are exclusive of HLA class I and II alleles. This is because B cells recognize antigens via B-cell receptors (BCR), known as membrane-bound immunoglobulins (Ig). Immunoglobulins consist of a constant fragment (Fc) region at the stalk and a variable (V) domain at the top. Given its functions in antigen binding, the V domain is responsible for the enormous theoretical diversity ( $10^{13-15}$ ) of the BCR repertoire [50,51]. Despite the high plasticity and diversity of the BCR repertoire, several lines of evidence demonstrated high frequencies of shared BCR clonotypes or elements in human BCR [52,53].

On this account, the conserved LBL epitopes were screened and identified using the ABCPred and SVMTriP prediction tools. These tools are considerably accurate in their predictions [54,55] and the results are analyzable. A total of 4238 peptides with thresholds of 0.5 and greater were obtained after eliminating the duplicated sequences. The antigenicity and toxicity predictions were performed as described earlier to exclude non-antigenic and toxic epitopes, leaving only 621 peptide sequences for further analysis. Subsequently, the number of epitopes was narrowed to 68 by retaining the peptides similar to these sequences or found in 50% or greater of the coronavirus strains (Table 8).

**Table 8.** Final selected LBL epitopes.

Peptide Sequence	Number of Matched Coronavirus Strains	Location in S Glycoprotein	Assigned Name
CVLGQSKRVDFCGKGY	25	1045–1060	LBL1
DKYFKNHTSPDVLGD	25	1166–1181	LBL2
DEDDSEPVLKGVKLHY	25	1270–1285	LBL3
AMQMAYFNGIGVTQN	25	899–914	LBL4
AGAALQIPFAMQMAYR	25	903–918	LBL5
FAMQMAYRFNGIGVTQ	25	911–926	LBL6
ASANLAATKMSECVLG	24	1033–1048	LBL7
ATKMSECVLGQSKRVD	24	1039–1054	LBL8
HGVVFLHVTVYVPAQEK	24	1071–1086	LBL9
HVTYVPAQEKNFITAP	24	1077–1092	LBL10
FVSGNCDVVGIVNNT	24	1134–1149	LBL11
VIGIVNNTVYDPLQPE	24	1142–1157	LBL12
HTSPDVLGDISGINA	24	1172–1187	LBL13
LGDISGINASVVNIQK	24	1179–1194	LBL14
GTTLDSKTQSLIVNN	24	120–135	LBL15
ESLIDLQELGKYEYI	24	1208–1223	LBL16
YVGYLQPRTFLLKYNE	24	279–294	LBL17
NENGTITDAVDCALDP	24	293–308	LBL18
AVDCALDPLSETKCTL	24	301–316	LBL19
DPLSETKCTLKSFTVE	24	307–322	LBL20
TVEKGIYQTSNFRVQP	24	320–335	LBL21
VQPTESIVRFPNITNL	24	333–348	LBL22
NDLCFTNVYADSFVIR	24	388–403	LBL23
PTKLNLCFTNVYADS	24	397–412	LBL24
VVLSFELLHAPATVCG	24	524–539	LBL25
FRSSVLHSTQDLFLPF	24	56–71	LBL26
TDAVRDPQTLEILDIT	24	586–601	LBL27
EILDITPCSFVGVSVI	24	596–611	LBL28
GVSVITPGTNTSNQVA	24	607–622	LBL29
HSTQDLFLPFFSNVTW	24	62–77	LBL30
YSTGSNVFQTRAGCLI	24	649–664	LBL31
TISVTTEILPVSMTKT	24	732–747	LBL32
TECSNLLLQYGSFTQ	24	760–775	LBL33
RALTGLIAVEQDKNTQE	24	778–793	LBL34
AVEQDKNTQEVFAQVK	24	784–799	LBL35
EMIAQYTSALLAGTIT	24	881–896	LBL36
AGTITSGWTFGAGAAL	24	892–907	LBL37
IGKIQDSLSTASALG	24	944–959	LBL38
FKCYGVSPKLNLCF	24	374–389	LBL39
FVTQRNFPYEQIITTD	23	1116–1131	LBL40
YEQYIKWPWYIWLGI	23	1219–1234	LBL41
PWYIWLGFIAGLIAIV	23	1226–1241	LBL42
EPLVDLPIGINITRFQ	23	237–252	LBL43
QTLALHRSYLTPGDS	23	239–254	LBL44
TRFQTLALHRSYLTP	23	249–264	LBL45
NQVAVLYQGVNCTEVP	23	606–621	LBL46
YQGVNCTEVPVAIHAD	23	612–627	LBL47
NNSIAIPTNFTISVTT	23	722–737	LBL48
RDLICAQKFNGLTVLP	23	860–875	LBL49
VFLVLLPLVSSQCVNL	22	16–31	LBL50
TGTGVLTESNKKFLPF	22	560–575	LBL51
NNSYECDIPIGAGICA	22	670–685	LBL52
SQSIIAYTMSLGAENS	22	702–717	LBL53
YTMSLGAENSVAYSNN	22	708–723	LBL54
GDCLGDIAARDLICAQ	22	851–866	LBL55
DIPIGAGICASYQTQT	21	663–678	LBL56

Table 8. Cont.

Peptide Sequence	Number of Matched Coronavirus Strains	Location in S Glycoprotein	Assigned Name
PFLMDLEGKQGNFKNL	20	187–202	LBL57
GWTAGAAAYVGYLQP	20	270–285	LBL58
HRSYLTPGDSSSGWTA	19	258–273	LBL59
YGVGHQPYRVVLSFE	19	501–516	LBL60
SYQTQTKSHRRARSVA	19	673–688	LBL61
TASALGKLQDVVNHNA	19	941–956	LBL62
KQLSSKFGAIVSVLND	19	964–979	LBL63
PVLPFNDGVYFASTEK	18	95–110	LBL64
PGQTGNIADYNYKLPD	17	412–427	LBL65
RKSNLKPFERDISTEI	17	470–485	LBL66
GSFCTQLKRALTGIIV	17	757–772	LBL67
LQSYGFRPTYGVGHQP	15	492–507	LBL68

3.5. Alignment and Assembly of the Identified HTL, CTL T, and LBL Epitopes

In this study, a total of 131 epitopes (12 CTL epitopes, 52 HTL epitopes, and 68 LBL epitopes) were identified. Generally, the selection criteria included (i) within the S1 or S2 region, (ii) conserved regions, and (iii) matching most HLA class I and II supertypes. The identified epitopes were aligned to the SARS-CoV-2-Wuhan-Hu-1 S glycoprotein to identify their positions (Supplementary Figure S4). They were then assembled and combined into two peptide sequences encompassing 39 and 34 amino acid residues, respectively (Table 9). Interestingly, the sequences of both assemblies, i.e., Epi1 and Epi2, corresponded to the epitopes located within the S1 region of the S glycoprotein. Epi1 was located at the N-terminal domain (NTD) (S<sub>256–294</sub>) while Epi2 was found in the RBD (S<sub>492–525</sub>). In addition, they were relatively conserved among the pangolin and bat coronaviral S glycoproteins (Table 3) except for the avian coronavirus strains. Notably, Epi1 was 66.7% (26/39) and 43.6% (17/39) similar to that of SARS-CoV and MERS-CoV, respectively (Supplementary Figure S3), suggesting that the conservation of Epi1 renders it a promising candidate of a broad-spectrum, cross-protective vaccine that potentially offers prophylactic protection against multiple coronavirus strains.

Table 9. Final assembled epitopes.

Combination of Peptides	Peptide Sequence	Peptide Location *	Peptide Length	Matched HLA Class I Supertype	Matched HLA Class II Supertype	Assigned Name
CTL3+ CTL4+ CTL6+ CTL7+ HTL50+ HTL42+ HTL30+ HTL31+ HTL43+ LBL59+ LBL58+ LBL17	SGWTAGA AAYVGYL QPRTFLK YNENGTIT DAVDCALD	256–294 (N-terminal domain)	39	A1, A2, A26, B8, B39, B58, B62	HLA-DPA1*01:03/DPB1*04:01; HLA-DPA1*01:03/DPB1*02:01; HLA-DPA1*02:01/DPB1*01:01; HLA-DPA1*02:01/DPB1*05:01; HLA-DPA1*03:01/DPB1*04:02; HLA-DQA1*01:01/DQB1*05:01; HLA-DQA1*01:02/DQB1*06:02; HLA-DQA1*04:01/DQB1*04:02; HLA-DQA1*05:01/DQB1*02:01; HLA-DQA1*05:01/DQB1*03:01; HLA-DRB1*01:01; HLA-DRB1*07:01; HLA-DRB1*09:01	Epi1
CTL1+ CTL10+ HTL51+ HTL25+ HTL14+ HTL22+ HTL23+ HTL15+ HTL16+ HTL45+ LBL60+ LBL68	LQSYGFQ PTNGVG YQPYRVV VLSFELL HAPATVC	492–525 (RBD)	34	A1, A2, A3, B7, B27, B58, B62	HLA-DPA1*01:03/DPB1*04:01; HLA-DPA1*01:03/DPB1*02:01; HLA-DPA1*02:01/DPB1*01:01; HLA-DPA1*02:01/DPB1*05:01; HLA-DPA1*03:01/DPB1*04:02; HLA-DQA1*01:01/DQB1*05:01; HLA-DQA1*03:01/DQB1*03:02; HLA-DQA1*05:01/DQB1*02:01; HLA-DRB1*01:01; HLA-DRB1*07:01; HLA-DRB1*09:01; HLA-DRB4*01:01	Epi2

\* Reference sequence: SARS-CoV-2-Wuhan-Hu-1 sequence.

Next, the allergenicity of Epi1 and Epi2 was determined by using AllergenFP v1.0. The results showed that Epi1 was a potential allergen whereas Epi2 was a non-allergen. Epi1 had the highest Tanimoto similarity index to a major allergen Pru av 1 (UniProtKB/Swiss-Prot ID: O24248) that causes birch pollinosis and oral allergy in patients allergic to cherry. Epi1 shared nine amino acids with that of the Pru av 1 peptide, albeit scattered throughout the latter's peptide sequence (Supplementary Figure S5). Nonetheless, the overlapping residues in the Epi1 and Pru av 1 peptides do not coincide with the known IgE binding regions of the Pru av 1 peptide, i.e., the P-loop region (44LEGDGGPGT52) [56,57], suggesting that the allergenic potential of Epi1 is mostly negligible. Furthermore, due to its relatively smaller molecular size, the 3D structure and physicochemical properties of Epi1 are not definitive, thus it is inconclusive to serve as an allergen after modification for vaccine development.

### 3.6. Population Coverage

Table 10 summarizes the findings of the population coverage of HLA classes of Epi1 and Epi2. The total world population coverage analysis showed that Epi1 and Epi2 matched with 75.53% and 81.06% with HLA class I. In terms of HLA class II, Epi1 showed a coverage of 99.88% whilst it was 99.74% for Epi2. When HLA class I and II were combined, the coverage for both epitopes was notably high, at 99.97% for Epi1 and 99.95% for Epi2. These results suggest that Epi1 and Epi2 are potential lead peptides for vaccine development, targeting broad global coverage and addressing the challenges posed by mutations in coronaviruses.

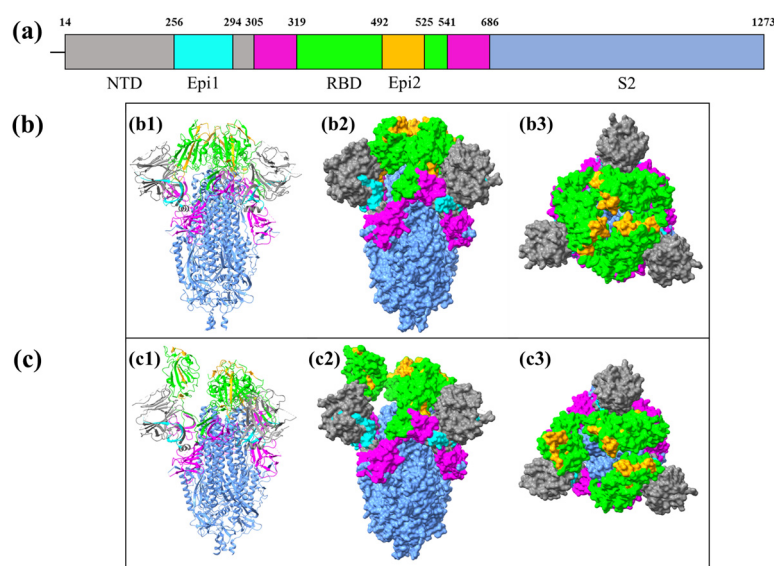
**Table 10.** Analysis of population coverage of HLA classes of Epi1 and Epi2.

Class	Epi1			Epi2		
	Coverage	Average Hit	PC90 *	Coverage	Average Hit	PC90 *
Class I	75.53%	1.09	0.41	81.06%	1.25	0.53
Class II	99.88%	4.27	3.04	99.74%	3.82	2.53
Combined	99.97%	5.36	3.79	99.95%	5.07	3.48

\* Minimum number of epitope hits/HLA combinations recognized by 90% of the population.

### 3.7. Identification of the Locations of the Conserved Epitopes in Coronaviral S Glycoprotein

As mentioned previously, Epi1 and Epi2 are located at the NTD (S<sub>256–294</sub>) and RBD (S<sub>492–525</sub>), respectively (Figure 3a). Figure 3 indicates the closed state and open state positions of Epi1 and Epi2 on the S glycoprotein.



**Figure 3.** Locations of the Assembled Epitopes in the S Glycoprotein. (a) Schematic diagram of SARS-CoV-2 S glycoprotein with different colors representing the S1 subunit (S<sub>14–685</sub>) (magenta) and



S2 subunit (S<sub>686–1273</sub>) (cornflower blue). **(b)** Ribbon and 3D structures of the closed state of SARS-CoV-2 S glycoprotein (PDB: 6VXX). **(b1)** The ribbon structure of closed state S glycoprotein. The locations of Epi1 and Epi2 were in cyan and orange colors, respectively. **(b2)** The orthogonal view of the closed state of S glycoprotein with Epi1 (cyan) and Epi2 (orange). **(b3)** The top-down view of the closed state of S glycoprotein displaying Epi1 (cyan) and Epi2 (orange). **(c)** Ribbon and 3D structures of the open state of SARS-CoV-2 S glycoprotein (PDB: 6VYB); **(c1)** The ribbon structure of the open state S glycoprotein and the locations of Epi1 and Epi2 in cyan and orange colors, respectively. **(c2)** The orthogonal view of the open state of S glycoprotein displaying Epi1 (cyan) and Epi2 (orange). **(c3)** The top-down view of the open state of S glycoprotein with Epi1 and Epi2 in cyan and orange colors, respectively.

The position of epitopes on an antigen contributes to its antigenicity and immunogenicity. The findings showed that Epi1 (cyan) was slightly embedded within the NTD (gray) and, therefore, was relatively less exposed than Epi2 (orange) located within the RBD domain (S<sub>319–541</sub>). This justifies the conservation of Epi1. Nonetheless, the conservation of epitopes is not solely determined by the exposure to immune cells or antibodies; it also depends on the functional importance of the epitopes. Mutations in conserved epitopes possibly disrupt key processes, such as viral attachment, entry, and immune evasion, thereby compromising the viral infectivity and replication in host cells [58–60]. In this light, conserved epitopes are important to ensure the structural and functional integrity of viral particles.

### 3.8. Molecular Docking of the Assembled Epitopes to TLR2 and TLR4 Receptors

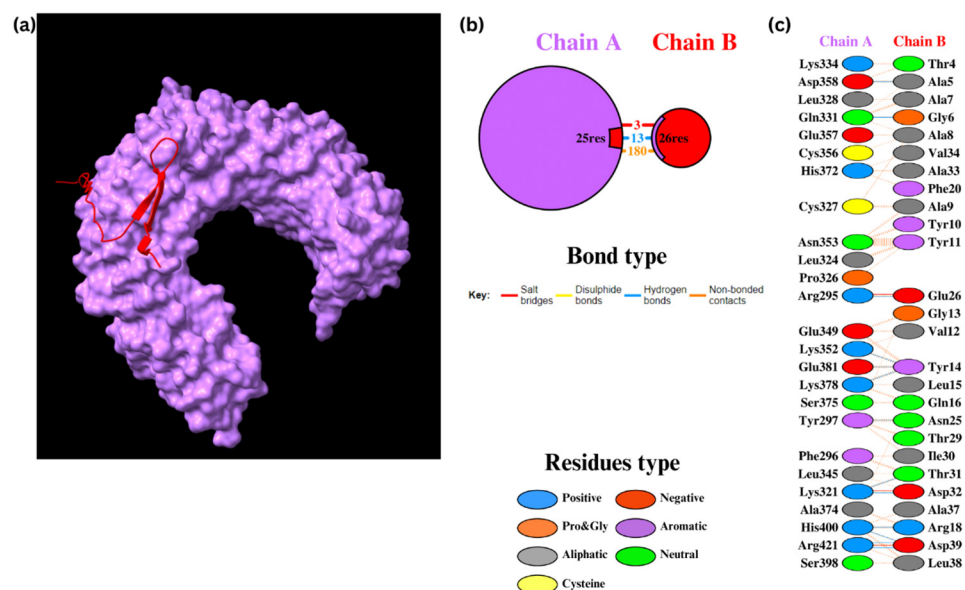
To mount a robust immune response, it is of utmost importance for a vaccine to interact with cell receptors. Molecular docking was therefore performed with Toll-like receptors, specifically TLR2 and TLR4, as these receptors are well recognized for their roles in interacting with viral structural proteins, which subsequently leads to inflammatory cytokine production [61,62].

#### 3.8.1. Docking of TLR2 with Epi1 and Epi2

The molecular docking analysis of Epi1 and Epi2 with TLR2 was performed using the HADDOCK 2.4 web server. For Epi1-TLR2, HADDOCK clustered 141 structures into 15 clusters, which represented 70% of the water-refined models (Figure 4). After refinement, 13 structures were clustered into one cluster resulting in 100% of the water-refined models. Epi1-TLR2 showed a strong binding affinity with a score of  $-76.6 \pm 5.2$ , in which a negative score demonstrates a good binding between the docked molecules (Table 11). The interaction was visualized using UCSF ChimeraX 1.8 and the bonding was analyzed using PDBsum. The results showed three salt bridges, 13 hydrogen bonds, and 180 non-bonded contacts.

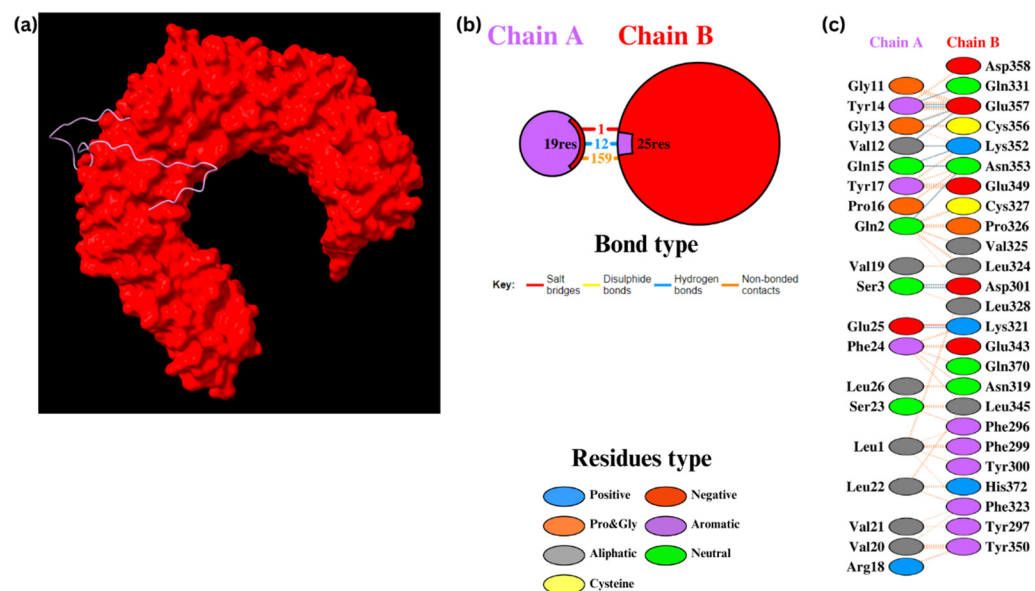
**Table 11.** HADDOCK 2.4 docking results for Epi1-TLR2 and Epi2-TLR2.

Scores	Epi1-TLR2	Epi2-TLR2
HADDOCK score	$-76.6 \pm 5.2$	$-82.4 \pm 6.5$
Cluster size	13	25
RMSD from the overall lowest-energy structure (Å)	$4.9 \pm 0.3$	$0.4 \pm 0.3$
Van der Waals energy (kcal mol <sup>-1</sup> )	$-63 \pm 4.5$	$-69.8 \pm 8.8$
Electrostatic energy (kcal mol <sup>-1</sup> )	$-255.2 \pm 32.3$	$-245.9 \pm 51.1$
Desolvation energy (kcal mol <sup>-1</sup> )	$-21.6 \pm 5.8$	$-37.3 \pm 4.5$
Restraints violation energy (kcal mol <sup>-1</sup> )	$589.9 \pm 65.3$	$738.6 \pm 17.6$
Buried surface area (Å <sup>2</sup> )	$2292.3 \pm 89.5$	$2271.0 \pm 76.4$
Z-Score	-1.5	-1.5



**Figure 4.** Docking of Epi1 with TLR2. (a) Epi1-TLR2 molecules. (b) The PDBsum results display the protein–protein interface between chain A (TLR2) and chain B (Epi1). (c) The PDBsum results display interacting amino acid residues in the interface.

Similarly, for Epi2-TLR2, HADDOCK clustered 145 structures into 12 clusters, which represented 72% of the water-refined models (Figure 5). Upon refinement, 25 structures were clustered into one cluster resulting in 100% of the water-refined models. Epi2-TLR2 showed a good binding affinity with a HADDOCK score of  $-82.4 \pm 6.5$  (Table 11). There were one salt bridge, 12 hydrogen bonds, and 159 non-bonded contacts between the molecules.



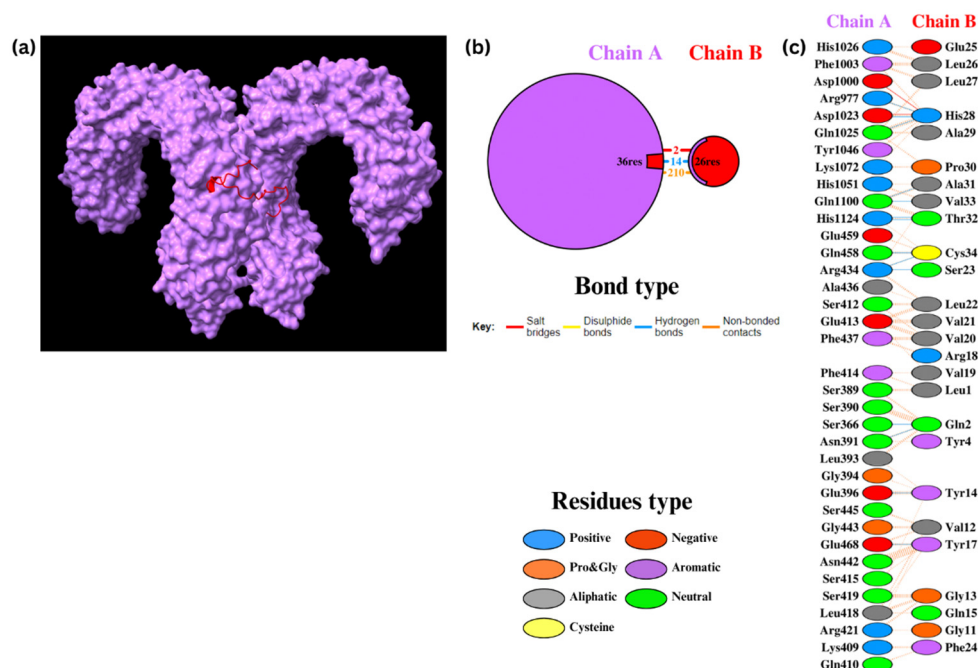
**Figure 5.** Docking of Epi2 with TLR2. (a) Epi2-TLR2 molecules. (b) The PDBsum results display the protein–protein interface between chain A (Epi2) and chain B (TLR2). (c) The PDBsum results demonstrate interactions between amino acid residues in the interface.

The binding affinities were further confirmed using the PRODIGY web server, which calculated the Gibbs free energy ( $\Delta G$ ) and the dissociation constants ( $K_d$ ) (Supplementary Table S2). Epi1-TLR2 and Epi2-TLR2 contained  $\Delta G$  of  $-14.8$  kcal/mol and  $-13.0$  kcal/mol,

respectively. The  $K_d$  were  $1.50 \times 10^{-11}$  M for Epi1 and  $2.80 \times 10^{-10}$  M for Epi2 at 25 °C. The relatively low  $\Delta G$  and  $K_d$  indicate strong, thermodynamically stable interactions.

### 3.8.2. Docking of TLR4 with Epi1 and Epi2

Epi1 and Epi2 were also docked to TLR4. For Epi1-TLR4, HADDOCK clustered 101 structures into 15 clusters, which represented 50% of the water-refined models (Figure 6). Subsequently, nine structures were refined and clustered into one cluster resulting in 100% of the water-refined models. The cluster had a HADDOCK score of  $24.2 \pm 19.4$  (Table 12) with 10 hydrogen bonds and 213 non-bonded contacts. The  $\Delta G$  was  $-17.1$  kcal/mol with a  $K_d$  of  $2.80 \times 10^{-13}$  M (Supplementary Table S2).



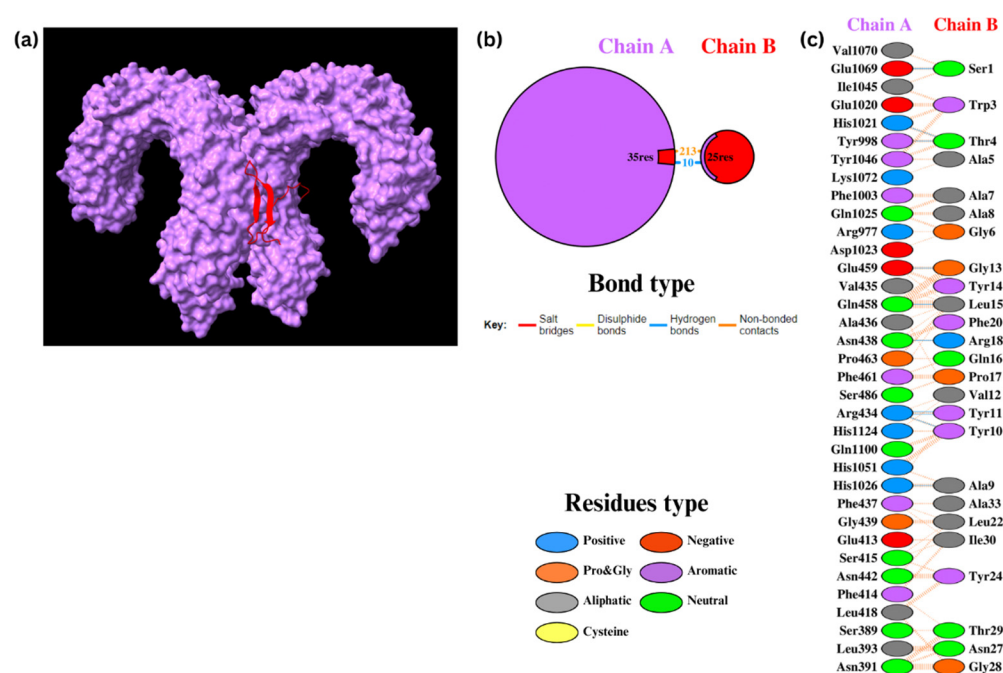
**Figure 6.** Docking of Epi1 to TLR4. (a) Epi1-TLR4 molecules. (b) The PDBsum results display the protein–protein interface between chain A (TLR4) and chain B (Epi1). (c) The PDBsum results display interactions between amino acid residues in the interface.

**Table 12.** HADDOCK 2.4 docking results for Epi1-TLR4 and Epi2-TLR4.

Scores	Epi1-TLR4	Epi2-TLR4
HADDOCK score	$24.2 \pm 19.4$	$27.7 \pm 4.6$
Cluster size	9	20
RMSD from the overall lowest energy structure (Å)	$0.5 \pm 0.3$	$8.4 \pm 0.0$
Van der Waals energy (kcal mol <sup>-1</sup> )	$-92.7 \pm 13.5$	$-101.1 \pm 6.5$
Electrostatic energy (kcal mol <sup>-1</sup> )	$-133.2 \pm 5.6$	$-125.3 \pm 11.0$
Desolvation energy (kcal mol <sup>-1</sup> )	$-56.0 \pm 4.1$	$-44.5 \pm 1.0$
Restraints violation energy (kcal mol <sup>-1</sup> )	$1995.3 \pm 111.3$	$1953.8 \pm 46.4$
Buried surface area (Å <sup>2</sup> )	$2719.0 \pm 164.8$	$2844.3 \pm 109.4$
Z-Score	-2.2	-1.8

For Epi2-TLR4, 113 structures were clustered in 17 clusters, which represents 56% of the water-refined models (Figure 7). After refinement, 20 structures were clustered into one cluster resulting in 100% of the water-refined models. A HADDOCK score of  $27.7 \pm 4.6$  (Table 12) was obtained with two salt bridges, 14 hydrogen bonds, and 210 non-bonded contacts. The  $\Delta G$  was  $-18.8$  kcal/mol with a  $K_d$  of  $1.70 \times 10^{-14}$  M (Supplementary Table S2), displaying a stronger binding affinity than that of Epi1. These

results indicate stable, thermodynamically favorable interactions between the assembled peptides and TLR4.



**Figure 7.** Docking of Epi2 to TLR4. (a) Epi2-TLR4 molecules. (b) The PDBsum results display the protein–protein interface between chain A (TLR4) and chain B (Epi2). (c) The PDBsum results display interactions between amino acid residues in the interface.

### 3.9. Immune Simulation

An in silico immune simulation was conducted using the C-IMMSIM server [29] to evaluate the immunostimulatory profile of Epi1 and Epi2. The simulation results of Epi1 and Epi2 were remarkably similar. The antibody titers shown in Figures 8a and 9a were zero, indicating lack of B cell responses. This aligns with Figures 8d,e and 9d,e, where immunoglobulin (Ig) isotype switching was absent. Typically, in the humoral response, proliferation of B lymphocytes into plasma cells involves an Ig class-switch recombination, leading to the production of Ig isotypes such as IgG and IgA that are crucial for neutralization of antigens [63]. However, the presence of only IgM-producing B cells implies that Ig isotype-switching did not occur as expected, particularly due to the lack of IL-4 and IL-21 (Figures 8i and 9i) [63]. A strong bias toward a cell-mediated immune response was, on the other hand, observed in the simulation (Figures 8f–h and 9f–h).

The antigen uptake and presentation of dendritic cells (DCs) and macrophages are shown in Figures 8b,c and 9b,c. In terms of Epi1 uptake and presentation on DCs, the antigen was preferably displayed on MHC class II to MHC class I. Interestingly, the presentation of Epi2 on DCs was predominantly mediated by MHC class I. Macrophages presented the epitopes mainly through MHC class II molecules.

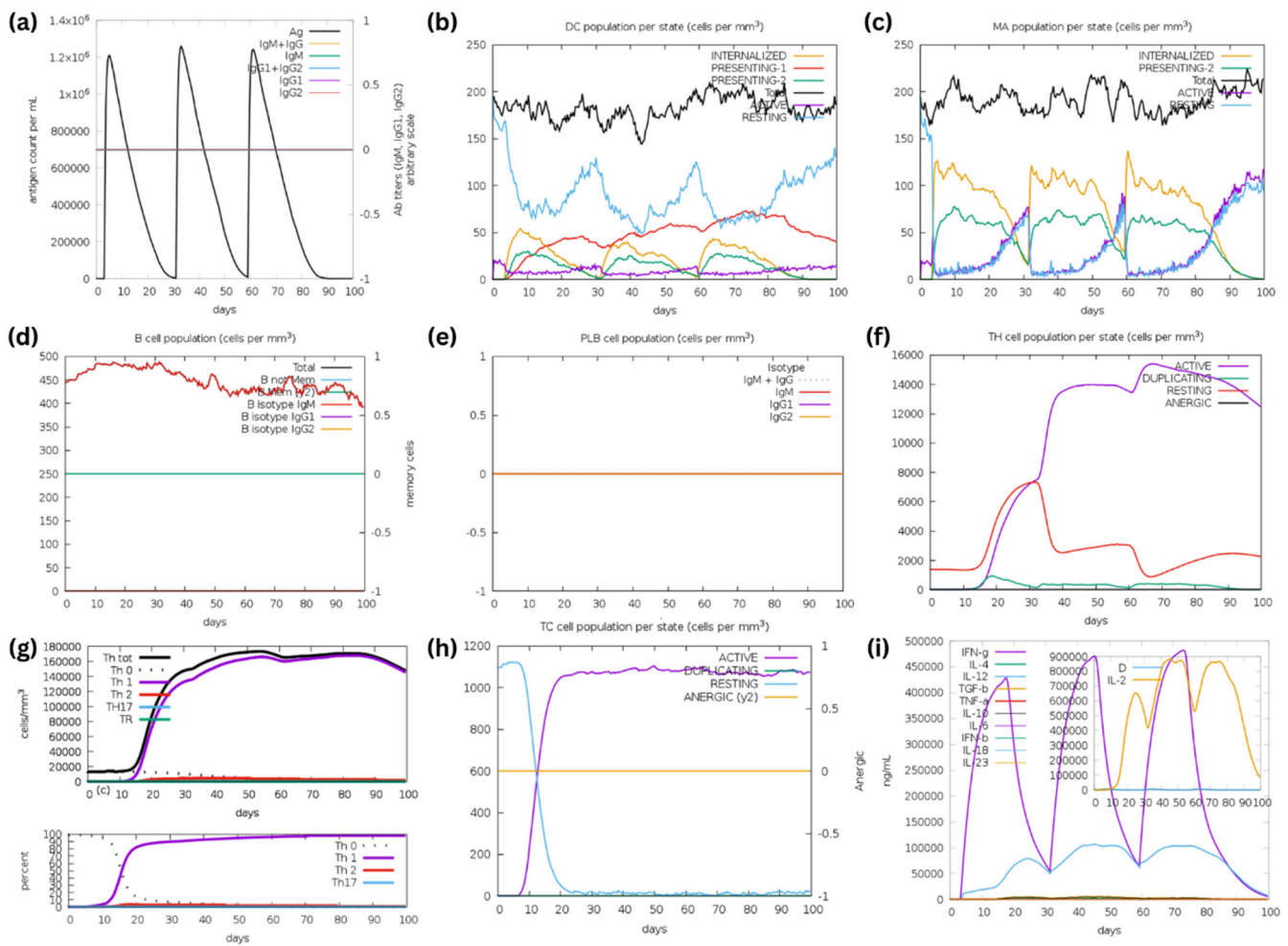
The hypothesis of Epi1 and Epi2 preferably stimulating cell-mediated immunity is supported by the findings in Figures 8f,g and 9f,g. The results implied a rapid and robust CD4+ T-helper cell response after the first immunization, with the T-helper cell counts exceeding 15,000 cells/mm<sup>3</sup> following the third injection. Notably, there was a clear bias toward Th1 differentiation (Figures 8g and 9g). This is in line with the findings of Figures 8h and 9h. Under the influence of IFN- $\gamma$ , IL-12 and IL-2 secreted by Th1, cytotoxic T (Tc) cells are stimulated and continue to proliferate into active Tc that are responsible for getting rid of infected cells in natural infections.

According to Figures 8i and 9i, the cytokine profile showed elevated IL-2, IFN- $\gamma$ , and IL-12 synthesis following the Epi1 and Epi2 vaccinations, indicating T-cell proliferation

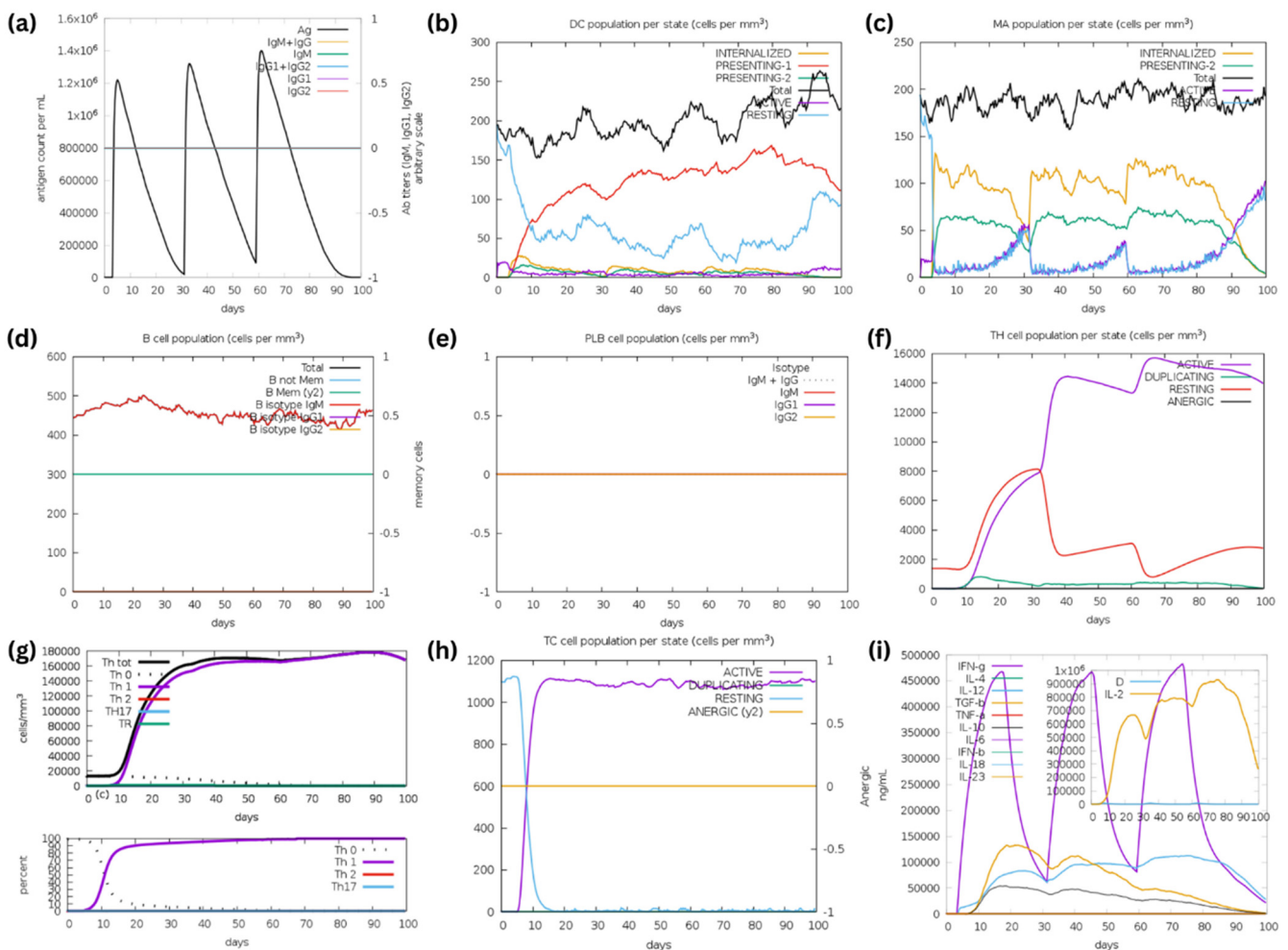


and expansion. Interestingly, the anti-inflammatory cytokines, i.e., TGF- $\beta$  and IL-10, were also observed at lower levels upon Epi2 immunization (Figure 9i), peaking at around 14,000 ng/mL and 5000 ng/mL, respectively. TGF- $\beta$  and IL-10 are known to be associated with the Th2 response with relatively lower TGF- $\beta$  and IL-10 synthesis, supporting the notion that Epi1 and Epi2 preferably mount Th1-biased immunity upon vaccination.

All in all, Epi1 and Epi2 elicited a predominantly cell-mediated immune response, characterized by strong Th1 and Tc cell activation. This phenomenon is seconded by robust production of IL-2, IFN- $\gamma$ , and IL-12.



**Figure 8.** The simulation of immunostimulatory potential of Epi1 (three injections) using the C-IMMSIM server. (a) Three doses of Epi1 were injected over a period of 3 months. (b) Antigen presentation on MHC class I and II molecules of dendritic cells (DC). The curves indicate various states of DC: active, resting, internalized and antigen presentation. (c) States of macrophages: total count, internalized, antigen presentation on MHC class II, active, and resting. (d) B lymphocytes: total count, memory cells, and Ig-isotypes, i.e., IgM, IgG1 and IgG2. (e) Plasma B lymphocytes: count, Ig-isotype (IgM, IgG1 and IgG2). (f) Total CD4+ T-helper lymphocyte count and states, i.e., active, resting, anergic and duplicating. (g) Counts of CD4+ T-helper lymphocyte subtypes. (h) Total CD8+ T-cytotoxic lymphocyte count and state. (i) IFN- $\gamma$ -, IL-12-, and IL-2-biased cytokine responses upon the peptide injections. D in the inset plot indicates the danger signal.



**Figure 9.** The simulation of immunostimulatory potential of Epi2 (three injections) using the C-IMMSIM server. (a) Three doses of Epi2 were injected over a period of 3 months. (b) Antigen presentation on MHC class I and II molecules of dendritic cells (DC). The curves indicate various states of DC: active, resting, internalized and antigen presentation. (c) States of macrophages: total count, internalized, antigen presentation on MHC class II, active, and resting. (d) B lymphocytes: total count, memory cells, and Ig-isotypes, i.e., IgM, IgG1 and IgG2. (e) Plasma B lymphocytes: count, Ig-isotype (IgM, IgG1 and IgG2). (f) Total CD4+ T-helper lymphocyte count and states, i.e., active, resting, anergic and duplicating. (g) Counts of CD4+ T-helper lymphocyte subtypes. (h) Total CD8+ T-cytotoxic lymphocyte count and state. (i) IFN- $\gamma$ -, IL-12-, and IL-2-biased cytokine responses upon the peptide injections. Some TGF- $\beta$  and IL-10 responses were also observed. D in the inset plot indicates the danger signal.

#### 4. Discussion

The 2019 SARS-CoV-2 pandemic revealed the unpreparedness of global healthcare systems to effectively respond to such a crisis, eventually leading to the breakdown of healthcare systems. The evolution and natural selection of coronaviruses are believed to contribute to the emergence of various VOCs with exceptional abilities to escape vaccine- and infection-induced immunity. The COVID-19 prophylactic vaccines are mainly based on the whole S glycoprotein subunit of SARS-CoV-2-Wuhan-1 due to its high antigenicity and immunogenicity [64–70]. However, the effectiveness of the COVID-19 vaccines is becoming less pronounced following the development of immune-evading variants due to the perpetual gene mutations [71–74]. In addition, the immune imprinting induced by immunization and previous infections also reduces the efficaciousness of the vaccines against newly emerged variants [75,76]. In this light, a conserved multi-epitope approach



has been adopted to develop pre-emptive vaccines against highly mutable coronaviruses by targeting the critical functional viral antigens. This strategy not only induces broad and long-lasting immune responses but it also prevents the need for a constant review of vaccine formulations due to viral mutations. To achieve this, epitope identification and characterization are entailed to generate epitope maps depicting their antibody specificities in silico prior to rigorous in vitro and in vivo empirical investigations [13–16,20–23,77–79].

Phylogenetically related zoonotic coronaviruses, including distantly related avian coronaviruses, were included in this study to identify and analyze the conserved regions. A significant genetic divergence was observed between avian and human coronaviruses compared to SARS-CoV-2. This divergence is especially notable when comparing avian coronaviruses, which belong to the *Gammacoronavirus* genus, with human coronaviruses and SARS-CoV-2, which belong to the *Alphacoronavirus* and *Betacoronavirus* genera, respectively. This observation is consistent with the phylogenetic data reported by Gilbert and Tengs (2021) [80]. It is noteworthy that none of the avian coronaviruses have been reported to infect humans to date. In this light, the prediction of conserved epitopes of coronaviruses prioritizes those and their close zoonotic counterparts causing diseases in humans (hCoVs, MERS-CoV, SARS-CoV, and SARS-CoV-2).

Initially, 12 HTL epitopes, 52 CTL epitopes, and 68 LBL epitopes were identified and the majority of them were conserved across the coronavirus strains. The avian coronaviruses, hCoVs, and MERS-CoV were distantly related to SARS-CoV, SARS-CoV-2, bat CoVs, and pangolin CoVs. The evolutionary convergence among those coronaviruses is likely due to the different natural and/or intermediate hosts [81]. In addition, it is noteworthy that the evolutionary convergence also results in the host-cell receptor variations as observed in hCoVs, among which the surface receptors responsible for viral adsorption are mainly surface peptidases and sialic acid-rich glycan-based receptors [82].

Human leukocyte antigen (HLA) alleles are among the most gene-dense and polymorphic regions in the human genome [83]. HLA molecules are responsible for antigen presentation to T-cell receptors (TcR) on CTL and HTL, and therefore can readily affect the vaccine-induced immune response [41,84–92]. In this light, it is important to retain antigenic epitopes that can interact and bind to HLA class I and class II molecules in vaccine design and development. Epi1 and Epi2 fulfill the characteristics of immunogenic vaccine candidates, given their multiple T-cell and B-cell epitopes. The total world population coverage analysis revealed a remarkably high class I coverage of above 75% and 81% for Epi1 and Epi2, respectively, whilst the class II coverage was greater than 99% for both Epi1 and Epi2. These results suggest the promising potential of these peptides as vaccine candidates with broad global MHC coverage, addressing the challenges posed by constant genome mutations in coronaviruses. Among the matched HLA class I supertypes, the A\*02 supertype is prevalently found in almost all human populations [93]. In regard to the HLA class II supertypes, more than 20 HLA class II supertypes were identified. Together with the LBL epitopes, Epi1 and Epi2 are expected to trigger cellular and humoral immune responses in vivo, thereby providing a more comprehensive protection against coronavirus infection [49,94].

Many of the identified epitope sequences overlapped one another; therefore, they were aligned and assembled into single peptide sequences. Two peptide assemblies, Epi1 and Epi2, consisting of HTL, CTL, and LBL epitopes, represented residues S256–294 (SGWTAGAAAYVGYLQPRITFLKYNENGTITDAVDCALD) and S492–525 (LQSYGFQPT-NGVGYQPYRVVLSFELLHAPATVC) of the S glycoprotein, respectively. Epi1 is located in the S1 region, particularly the NTD; Epi2, on the other hand, is found in the RBD region. Essentially, the S glycoprotein of SARS-CoV-2, one of the structural components of the virus, was reported to interact with TLR4 and potentially TLR2, thus leading to inflammatory responses [95–99]. In humans, TLR4 are predominantly expressed in cells of myeloid origin such as macrophages, immature dendritic cells, monocytes, and granulocytes [100]. TLR4 signaling is initiated when the receptor binds to a ligand, leading to homodimerization and recruitment of Toll/interleukin-1 receptor-like (TIR)-domain-containing adapter

molecules. This activates two major pathways: (i) the Myeloid differentiation primary response 8-dependent (MyD88-dependent) pathway, responsible for early nuclear factor kappa- $\beta$  (NF- $\kappa$ B) activation and pro-inflammatory cytokine release, and (ii) the TIR domain-containing adaptor inducing IFN- $\beta$ -TRIF-related adaptor molecule pathway (TRIF-TRAM pathway), which induces type-I interferon production and TNF- $\alpha$  secretion for late NF- $\kappa$ B activation [100]. On the other hand, human TLR2 is mainly expressed on myelomonocytic cell lines and specific blood cells with CD14<sup>+</sup> monocytes showing the highest expression, followed by CD15<sup>+</sup> granulocytes [101]. TLR2 signaling is initiated through ligand-induced dimerization with other TLRs, such as TLR1, TLR4, TLR6, and TLR10 [102–105], enabling the recognition of diverse microbial products and viral proteins [106–109]. This interaction subsequently activates MyD88-dependent signaling, leading to phosphorylation events involving interleukin-1 receptor-associated kinase (IRAK) proteins and receptor-associated factor 6 (TRAF6), which in turn activates NF- $\kappa$ B and mitogen-activated protein kinase (MAPK) pathways [110]. These pathways promote the production of pro-inflammatory cytokines and modulate cell proliferation and survival [111,112]. In this light, Epi1 and Epi2 were docked to TLR2 and TLR4 to estimate their interactions and binding strength. The docking results showed relatively good HADDOCK scores for both Epi1-TLR2 and Epi2-TLR2, indicating the potential of these peptides to activate the TLR-2-mediated innate immune response. Similarly, the HADDOCK scores of Epi1-TLR4 and Epi2-TLR4 were reasonable, implying the activation of TLR4 signaling upon binding to Epi1 and Epi2. These observations were further validated in the PRODIGY and PDBsum analyses, in which stable and favorable interactions were formed between the complexes. All in all, the standard HADDOCK 2.4 protocol effectively combines rigid docking with molecular dynamics simulations.

Epi1 and Epi2 were shown to predominantly elicit a cell-mediated immune response, characterized by strong Th1 and Tc cell activation with elevated levels of IL-2, IFN- $\gamma$ , and IL-12 after *in silico* immunization. Notably, antigen presentation by macrophages and dendritic cells on the MHC class II molecules is responsible for stimulating CD4<sup>+</sup> T-helper cells. Concomitantly, dendritic cells were also demonstrated to display epitopes of Epi2 on MHC class I molecules, thereby promoting the activation of cytotoxic T cells [113]. The activation of CD4<sup>+</sup> T cells also results in the elevation of IL-2, which, in turn, promotes Th1 differentiation [114,115]. In addition, IL-12 secreted by dendritic cells and macrophages, also boosts Th1 differentiation for IFN- $\gamma$  production [115,116], a key pro-inflammatory cytokine essential for orchestrating cell-mediated immunity [117]. Interestingly, the *in silico* simulation results coincided with several *in vitro* and *in vivo* reports. For instance, Meyer et al. (2023) reported that the S269–277 epitope, which is similar to that of Epi1, could induce a high magnitude of CTL immune response after *in vitro* stimulation [19]. An *in vivo* study confirmed that a peptide segment identical to that of Epi1 (S265–279 or YYVGYLQPRTFLLKY) could induce a robust antigen-specific IFN- $\gamma$ -producing CTL response [17]; meanwhile, another *in silico* study identified the SGWTAGAAAYYV motif found in Epi1 as the immunodominant site for T-cell and humoral responses [12]. Collectively, Epi1 is a robust candidate for the development of multi-epitope vaccines. Although Epi2 (S492–525) and its immunostimulatory roles have not been reported elsewhere, its PYRVVLSF motif was hypothesized to induce adaptive immunity [11].

Given the conservation of Epi1 and Epi2, they hold promise as lead antigens in universal multi-epitope vaccine development, particularly in fighting the upcoming mutants. This helps address issues concerning the constant gene mutations and immune evasion seen in coronaviruses. Epi2 consists of most of the important residues required to form tight binding with ACE2 receptors [118–120]. It also encompasses well-known mutation sites found in the currently circulating Omicron variant, i.e., N501 and Y505. The N501Y mutation can lower neutralizing antibody binding *in vitro* [121], while the Y505H mutation reduces viral protein stability, affects viral infectivity, and promotes immune evasion [122,123]. Given the importance of the mutations, including them in a vaccine formulation is likely to add to

the relevance of the multi-epitope vaccine with the circulating coronavirus variants, hence greater immune protection.

Incorporating multiple epitopes in a vaccine formulation can offer broader and more durable protection against a wider range of viral variants. The multi-epitope sequences identified in this study shed light on the ongoing development and applications of coronavirus vaccines. The immune simulation analysis indicates that the medium length assembled peptides have limitations in stimulating humoral immune responses, as evidenced by the results (Figures 8 and 9). This could be attributed to the C-IMMSIM web server's restriction to MHC class II DRB alleles selection, having no alleles like DPA, DPB, DQA, or DQB in the MHC class II selection, which was found to be more prevalent in those assembled peptides. The interaction of MHC class II and peptides are crucial. The binding between peptides and the MHC molecules may promote humoral responses in vivo through both (i) direct mechanisms involving antigen presentation by APCs or (ii) indirect mechanisms where the interaction of the MHC receptor of CD4+ T cells interacts with B-cell MHC class II molecules, and/or (iii) through the release of cytokines by CD4+ T cells [63]. While stimulation of CD4+ and CD8+ T cells is likely to promote humoral responses in vivo, this needs to be validated through further in vivo studies. Additionally, in order to effectively deliver the assembled peptide and further improve its immunogenicity in vivo, the peptide can also be conjugated with virus-like particles (VLPs) or nanoparticle-based delivery systems to further enhance their abilities in inducing humoral and cellular responses [124,125]. Furthermore, the epitopes can be developed into multivalent vaccines consisting of promising flu antigens such as nucleoprotein (NP) of influenza A virus (IAV), which assembles into virus-like particles (VLP) for vaccine delivery [30,126]. To strengthen the multivalency of the vaccine, highly conserved matrix 2 ectodomain protein (M2e) of IAV can be added into the vaccine formulation. The IAV M2e is known for conferring partial protection in animal models against IAV [127,128]. Collectively, the aforementioned prospective applications highlight the versatility and the potential of these multi-epitope peptides in curbing coronavirus infections.

Although these predictive models allow us to screen epitopes quickly, they may not fully capture the complexities of antigen processing, epitope conformations, or interplays among immune cells and signaling pathways. This limited scope means in silico predictions might not fully correlate with the in vitro or in vivo immunogenicity and effectiveness of vaccine candidates. For instance, the C-IMMSIM predictions applied in this study may restrict certain MHC allele selections (e.g., DRB alleles), potentially affecting the overall humoral response prediction. Thus, in vitro and in vivo studies are warranted to further validate the potential of Epi1 and Epi2 as vaccine candidates. All in all, this study revealed evolutionarily conserved regions within SARS-CoV-2, SARS-CoV, and certain animal coronaviruses, while highlighting the genetic divergence observed in MERS-CoV and hCoVs as implied by their minimal sequence similarities in the S1 subunit.

## 5. Conclusions

In conclusion, evolutionarily conserved epitopes are present among animal and human coronaviral S glycoproteins. Overall, 132 candidates representing HTL, CTL and LBL epitopes with relatively low evolutionary divergence were identified. They were screened and filtered into two final peptide assemblies: Epi1 is composed of four HLA class I, five HLA class II, and three LBL epitopes; meanwhile, Epi2 consists of two HLA class I, eight HLA class II, and two LBL epitopes. Both peptides are located within the S1 subunit of the coronaviral S protein and demonstrate high population coverage and conservation. Additionally, they are expected to exhibit robust immunomodulatory roles, particularly cellular immunity via TLR2- and TLR4-mediated immune responses. Notably, Epi1 also contains immunodominant CTL epitopes, which adds to its potential as a vaccine candidate. Collectively, the conserved epitopes provide a robust foundation for universal vaccine development with extraordinary abilities to stimulate broad-spectrum immunity to mitigate the impact of coronaviral infections.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/biomedicines12112530/s1>, Table S1: Final selected HTL epitopes; Table S2: Complete PRODIGY analysis for all of the docked complexes; Figure S1: Animal coronaviruses alignment result with SARS-CoV-2-Wuhan-Hu-1 strain as the reference sequence; Figure S2: Human coronaviruses alignment result with SARS-CoV-2-Wuhan-Hu-1 strain as the reference sequence; Figure S3: Alignment result of SARS-CoV, MERS-CoV, and all SARS-CoV-2 variants with SARS-CoV-2-Wuhan-Hu-1 strain as the reference sequence; Figure S4: Alignment result of assembled epitopes, Sepi1 and Sepi2, to SARS-CoV-2-Wuhan-Hu-1 sequence; Figure S5: Alignment result of Sepi1 to Pru av 1 allergen peptide sequence.

**Author Contributions:** Conceptualization, Y.C.O., B.A.T. and W.B.Y.; methodology and investigation, Y.C.O. and W.B.Y.; data analysis and validation, Y.C.O., B.A.T. and W.B.Y.; writing—original draft preparation, Y.C.O.; writing—review and editing, Y.C.O., B.A.T. and W.B.Y.; supervision, B.A.T. and W.B.Y.; project administration, W.B.Y.; funding acquisition, W.B.Y. All authors have read and agreed to the published version of the manuscript.

**Funding:** This study was supported by the Fundamental Research Grant Scheme (FRGS/1/2021/STG03/UKM/02/2) from the Ministry of Higher Education, Malaysia.

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

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author.

**Acknowledgments:** The postgraduate student and partial publication fee were sponsored by the United States Agency for International Development (USAID) through SEAHOHUN 2024 Scholarship Program. The contents are the responsibility of the author(s) and do not necessarily reflect the views of USAID or the United States Government.

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

## References

1. Pormohammad, A.; Ghorbani, S.; Khatami, A.; Farzi, R.; Baradaran, B.; Turner, D.L.; Turner, R.J.; Bahr, N.C.; Idrovo, J.P. Comparison of confirmed COVID-19 with SARS and MERS cases—Clinical characteristics, laboratory findings, radiographic signs and outcomes: A systematic review and meta-analysis. *Rev. Med. Virol.* **2020**, *30*, e2112. [[CrossRef](#)]
2. Hui, D.S.; Perlman, S.; Zumla, A. Spread of MERS to South Korea and China. *Lancet Respir. Med.* **2015**, *3*, 509–510. [[CrossRef](#)] [[PubMed](#)]
3. WHO Director-General’s Opening Remarks at the Media Briefing on COVID-19—11 March 2020. Available online: <https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020> (accessed on 14 August 2024).
4. COVID-19 Cases | WHO COVID-19 Dashboard. Available online: <https://data.who.int/dashboards/covid19/cases> (accessed on 14 August 2024).
5. Andrews, N.; Stowe, J.; Kirsebom, F.; Toffa, S.; Rickeard, T.; Gallagher, E.; Gower, C.; Kall, M.; Groves, N.; O’Connell, A.-M.; et al. COVID-19 Vaccine Effectiveness against the Omicron (B.1.1.529) Variant. *N. Engl. J. Med.* **2022**, *386*, 1532–1546. [[CrossRef](#)]
6. Greaney, A.J.; Starr, T.N.; Gilchuk, P.; Zost, S.J.; Binshtein, E.; Loes, A.N.; Hilton, S.K.; Huddleston, J.; Eguia, R.; Crawford, K.H.D.; et al. Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding Domain that Escape Antibody Recognition. *Cell Host Microbe* **2021**, *29*, 44–57.e9. [[CrossRef](#)] [[PubMed](#)]
7. Harvey, W.T.; Carabelli, A.M.; Jackson, B.; Gupta, R.K.; Thomson, E.C.; Harrison, E.M.; Ludden, C.; Reeve, R.; Rambaut, A.; Peacock, S.J.; et al. SARS-CoV-2 variants, spike mutations and immune escape. *Nat. Rev. Microbiol.* **2021**, *19*, 409–424. [[CrossRef](#)] [[PubMed](#)]
8. Malik, J.A.; Ahmed, S.; Mir, A.; Shinde, M.; Bender, O.; Alshammari, F.; Ansari, M.; Anwar, S. The SARS-CoV-2 mutations versus vaccine effectiveness: New opportunities to new challenges. *J. Infect. Public Health* **2022**, *15*, 228–240. [[CrossRef](#)]
9. McLean, G.; Kamil, J.; Lee, B.; Moore, P.; Schulz, T.F.; Muik, A.; Sahin, U.; Türeci, Ö.; Pather, S. The Impact of Evolving SARS-CoV-2 Mutations and Variants on COVID-19 Vaccines. *mBio* **2022**, *13*, e02979–21. [[CrossRef](#)]
10. Prakash, S.; Srivastava, R.; Coulon, P.-G.; Dhanushkodi, N.R.; Chentoufi, A.A.; Tifrea, D.F.; Edwards, R.A.; Figueroa, C.J.; Schubl, S.D.; Hsieh, L.; et al. Genome-Wide B Cell, CD4+, and CD8+ T Cell Epitopes That Are Highly Conserved between Human and Animal Coronaviruses, Identified from SARS-CoV-2 as Targets for Preemptive Pan-Coronavirus Vaccines. *J. Immunol.* **2021**, *206*, 2566–2582. [[CrossRef](#)]
11. Bagherzadeh, M.A.; Izadi, M.; Baesi, K.; Jahromi, M.A.M.; Pirestani, M. Considering epitopes conservity in targeting SARS-CoV-2 mutations in variants: A novel immunoinformatics approach to vaccine design. *Sci. Rep.* **2022**, *12*, 14017. [[CrossRef](#)]



12. Jaiswal, V.; Lee, H.J. Conservation and Evolution of Antigenic Determinants of SARS-CoV-2: An Insight for Immune Escape and Vaccine Design. *Front. Immunol.* **2022**, *13*, 832106. [[CrossRef](#)]
13. Ahmed, S.F.; Quadeer, A.A.; McKay, M.R. Preliminary Identification of Potential Vaccine Targets for the COVID-19 Coronavirus (SARS-CoV-2) Based on SARS-CoV Immunological Studies. *Viruses* **2020**, *12*, 254. [[CrossRef](#)] [[PubMed](#)]
14. Ismail, S.; Ahmad, S.; Azam, S.S. Immunoinformatics characterization of SARS-CoV-2 spike glycoprotein for prioritization of epitope based multivalent peptide vaccine. *J. Mol. Liq.* **2020**, *314*, 113612. [[CrossRef](#)] [[PubMed](#)]
15. Maleki, A.; Russo, G.; Parasiliti Palumbo, G.A.; Pappalardo, F. In silico design of recombinant multi-epitope vaccine against influenza A virus. *BMC Bioinform.* **2022**, *22*, 617. [[CrossRef](#)] [[PubMed](#)]
16. Russo, G.; Crispino, E.; Maleki, A.; Di Salvatore, V.; Stanco, F.; Pappalardo, F. Beyond the state of the art of reverse vaccinology: Predicting vaccine efficacy with the universal immune system simulator for influenza. *BMC Bioinform.* **2023**, *24*, 231. [[CrossRef](#)] [[PubMed](#)]
17. Jiang, S.; Wu, S.; Zhao, G.; He, Y.; Guo, X.; Zhang, Z.; Hou, J.; Ding, Y.; Cheng, A.; Wang, B. Identification of a promiscuous conserved CTL epitope within the SARS-CoV-2 spike protein. *Emerg. Microbes Infect.* **2022**, *11*, 730. [[CrossRef](#)] [[PubMed](#)]
18. Smith, T.R.F.; Patel, A.; Ramos, S.; Elwood, D.; Zhu, X.; Yan, J.; Gary, 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](#)]
19. Meyer, S.; Blaas, I.; Bollineni, R.C.; Delic-Sarac, M.; Tran, T.T.; Knetter, C.; Dai, K.-Z.; Madssen, T.S.; Vaage, J.T.; Gustavsen, A.; et al. Prevalent and immunodominant CD8 T cell epitopes are conserved in SARS-CoV-2 variants. *Cell Rep.* **2023**, *42*, 111995. [[CrossRef](#)]
20. Mishra, N.; Huang, X.; Joshi, S.; Guo, C.; Ng, J.; Thakkar, R.; Wu, Y.; Dong, X.; Li, Q.; Pinapati, R.S.; et al. Immunoreactive peptide maps of SARS-CoV-2. *Commun. Biol.* **2021**, *4*, 225. [[CrossRef](#)]
21. Wang, H.; Wu, X.; Zhang, X.; Hou, X.; Liang, T.; Wang, D.; Teng, F.; Dai, J.; Duan, H.; Guo, S.; et al. SARS-CoV-2 Proteome Microarray for Mapping COVID-19 Antibody Interactions at Amino Acid Resolution. *ACS Cent. Sci.* **2020**, *6*, 2238–2249. [[CrossRef](#)]
22. Sikora, M.; von Bülow, S.; Blanc, F.E.C.; Gecht, M.; Covino, R.; Hummer, G. Computational epitope map of SARS-CoV-2 spike protein. *PLoS Comput. Biol.* **2021**, *17*, e1008790. [[CrossRef](#)]
23. Schwarz, T.; Heiss, K.; Mahendran, Y.; Casilag, F.; Kurth, F.; Sander, L.E.; Wendtner, C.-M.; Hoehstetter, M.A.; Müller, 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](#)] [[PubMed](#)]
24. Paul, S.; Arlehamn, C.S.L.; Scriba, T.J.; Dillon, M.B.C.; Oseroff, C.; Hinz, D.; McKinney, D.M.; Pro, S.C.; Sidney, J.; Peters, B.; et al. Development and validation of a broad scheme for prediction of HLA class II restricted T cell epitopes. *J. Immunol. Methods* **2015**, *422*, 28–34. [[CrossRef](#)] [[PubMed](#)]
25. Patel, M.C.; Shirey, K.A.; Pletneva, L.M.; Boukhvalova, M.S.; Garzino-Demo, A.; Vogel, S.N.; Blanco, J.C. Novel drugs targeting Toll-like receptors for antiviral therapy. *Future Virol.* **2014**, *9*, 811–829. [[CrossRef](#)] [[PubMed](#)]
26. Chen, J.; Ng, M.M.-L.; Chu, J.J.H. Activation of TLR2 and TLR6 by Dengue NS1 Protein and Its Implications in the Immunopathogenesis of Dengue Virus Infection. *PLoS Pathog.* **2015**, *11*, e1005053. [[CrossRef](#)]
27. Vangone, A.; Bonvin, A.M. Contacts-based prediction of binding affinity in protein–protein complexes. *eLife* **2015**, *4*, e07454. [[CrossRef](#)]
28. Xue, L.C.; Rodrigues, J.P.; Kastritis, P.L.; Bonvin, A.M.; Vangone, A. PRODIGY: A web server for predicting the binding affinity of protein–protein complexes. *Bioinformatics* **2016**, *32*, 3676–3678. [[CrossRef](#)]
29. Rapin, N.; Lund, O.; Bernaschi, M.; Castiglione, F. Computational Immunology Meets Bioinformatics: The Use of Prediction Tools for Molecular Binding in the Simulation of the Immune System. *PLoS ONE* **2010**, *5*, e9862. [[CrossRef](#)]
30. Chenavas, S.; Estrozi, L.F.; Slama-Schwok, A.; Delmas, B.; Primo, C.D.; Baudin, F.; Li, X.; Crépin, T.; Ruigrok, R.W.H. Monomeric Nucleoprotein of Influenza A Virus. *PLoS Pathog.* **2013**, *9*, e1003275. [[CrossRef](#)]
31. Kirchdoerfer, R.N.; Cottrell, C.A.; Wang, N.; Pallesen, J.; Yassine, H.M.; Turner, H.L.; Corbett, K.S.; Graham, B.S.; McLellan, J.S.; Ward, A.B. Pre-fusion structure of a human coronavirus spike protein. *Nature* **2016**, *531*, 118–121. [[CrossRef](#)]
32. Li, Z.; Tomlinson, A.C.A.; Wong, A.H.M.; Zhou, D.; Desforges, M.; Talbot, P.J.; Benlekbir, S.; Rubinstein, J.L.; Rini, J.M. The human coronavirus HCoV-229E S-protein structure and receptor binding. *eLife* **2019**, *8*, e51230. [[CrossRef](#)]
33. Wang, C.; Hesketh, E.L.; Shamorkina, T.M.; Li, W.; Franken, P.J.; Drabek, D.; van Haperen, R.; Townend, S.; van Kuppeveld, F.J.M.; Grosveld, F.; et al. Antigenic structure of the human coronavirus OC43 spike reveals exposed and occluded neutralizing epitopes. *Nat. Commun.* **2022**, *13*, 2921. [[CrossRef](#)] [[PubMed](#)]
34. Huang, Y.; Yang, C.; Xu, X.; Xu, W.; Liu, S. Structural and functional properties of SARS-CoV-2 spike protein: Potential antiviral drug development for COVID-19. *Acta Pharmacol. Sin.* **2020**, *41*, 1141–1149. [[CrossRef](#)] [[PubMed](#)]
35. Xia, X. Domains and Functions of Spike Protein in SARS-CoV-2 in the Context of Vaccine Design. *Viruses* **2021**, *13*, 109. [[CrossRef](#)]
36. Zheng, Z.; Monteil, V.M.; Maurer-Stroh, S.; Yew, C.W.; Leong, C.; Mohd-Ismail, N.K.; Arularasu, S.C.; Chow, V.T.K.; Lin, R.T.P.; Mirazimi, A.; et al. Monoclonal antibodies for the S2 subunit of spike of SARS-CoV-1 cross-react with the newly-emerged SARS-CoV-2. *Eurosurveillance* **2020**, *25*, 19–28. [[CrossRef](#)]
37. Boni, M.F.; Lemey, P.; Jiang, X.; Lam, T.T.-Y.; Perry, B.W.; Castoe, T.A.; Rambaut, A.; Robertson, D.L. Evolutionary origins of the SARS-CoV-2 sarbecovirus lineage responsible for the COVID-19 pandemic. *Nat. Microbiol.* **2020**, *5*, 1408–1417. [[CrossRef](#)]
38. Li, X.; Giorgi, E.E.; Marichannegowda, M.H.; Foley, B.; Xiao, C.; Kong, X.-P.; Chen, Y.; Gnanakaran, S.; Korber, B.; Gao, F. Emergence of SARS-CoV-2 through recombination and strong purifying selection. *Sci. Adv.* **2020**, *6*, eabb9153. [[CrossRef](#)]

39. Sajini, A.A.; Alkayyal, A.A.; Mubarak, F.A. The Recombination Potential between SARS-CoV-2 and MERS-CoV from Cross-Species Spill-over Infections. *J. Epidemiol. Glob. Health* **2021**, *11*, 155–159. [CrossRef] [PubMed]
40. 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]
41. Larsen, M.V.; Lundegaard, C.; Lamberth, K.; Buus, S.; Brunak, S.; Lund, O.; Nielsen, M. An integrative approach to CTL epitope prediction: A combined algorithm integrating MHC class I binding, TAP transport efficiency, and proteasomal cleavage predictions. *Eur. J. Immunol.* **2005**, *35*, 2295–2303. [CrossRef]
42. Sidney, J.; Grey, H.M.; Kubo, R.T.; Sette, A. Practical, biochemical and evolutionary implications of the discovery of HLA class I supermotifs. *Immunol. Today* **1996**, *17*, 261–266. [CrossRef]
43. Sidney, J.; Peters, B.; Frahm, N.; Brander, C.; Sette, A. HLA class I supertypes: A revised and updated classification. *BMC Immunol.* **2008**, *9*, 1. [CrossRef] [PubMed]
44. dos Santos Francisco, R.; Buhler, S.; Nunes, J.M.; Bitarello, B.D.; França, G.S.; Meyer, D.; Sanchez-Mazas, A. HLA supertype variation across populations: New insights into the role of natural selection in the evolution of HLA-A and HLA-B polymorphisms. *Immunogenetics* **2015**, *67*, 651–663. [CrossRef] [PubMed]
45. Andreatta, M.; Trolle, T.; Yan, Z.; Greenbaum, J.A.; Peters, B.; Nielsen, M. An automated benchmarking platform for MHC class II binding prediction methods. *Bioinformatics* **2018**, *34*, 1522–1528. [CrossRef] [PubMed]
46. Wang, P.; Sidney, J.; Dow, C.; Mothé, B.; Sette, A.; Peters, B. A Systematic Assessment of MHC Class II Peptide Binding Predictions and Evaluation of a Consensus Approach. *PLoS Comput. Biol.* **2008**, *4*, e1000048. [CrossRef] [PubMed]
47. Greenbaum, J.; Sidney, J.; Chung, J.; Brander, C.; Peters, B.; Sette, A. Functional classification of class II human leukocyte antigen (HLA) molecules reveals seven different supertypes and a surprising degree of repertoire sharing across supertypes. *Immunogenetics* **2011**, *63*, 325–335. [CrossRef]
48. Regenmortel, M.H.V. What Is a B-Cell Epitope? In *Epitope Mapping Protocols*; Schutkowski, M., Reineke, U., Eds.; Humana Press: Totowa, NJ, USA, 2009; pp. 3–20.
49. Sanchez-Trincado, J.L.; Gomez-Perosanz, M.; Reche, P.A. Fundamentals and Methods for T- and B-Cell Epitope Prediction. *J. Immunol. Res.* **2017**, *2017*, 2680160. [CrossRef]
50. Chaudhary, N.; Wesemann, D.R. Analyzing Immunoglobulin Repertoires. *Front Immunol.* **2018**, *9*. [CrossRef]
51. Raybould, M.I.J.; Rees, A.R.; Deane, C.M. Current strategies for detecting functional convergence across B-cell receptor repertoires. *MAbs* **2021**, *13*, 1996732. [CrossRef]
52. Briney, B.; Inderbitzin, A.; Joyce, C.; Burton, D.R. Commonality despite exceptional diversity in the baseline human antibody repertoire. *Nature* **2019**, *566*, 393–397. [CrossRef]
53. Soto, C.; Bombardi, R.G.; Branchizio, A.; Kose, N.; Matta, P.; Sevy, A.M.; Sinkovits, R.S.; Gilchuk, P.; Finn, J.A.; Crowe, J.E. High frequency of shared clonotypes in human B cell receptor repertoires. *Nature* **2019**, *566*, 398–402. [CrossRef]
54. Galanis, K.A.; Nastou, K.C.; Papandreou, N.C.; Petichakis, G.N.; Iconomidou, V.A. Linear B-Cell Epitope Prediction: A Performance Review of Currently Available Methods. 2019. Available online: <https://www.biorxiv.org/content/10.1101/833418v1> (accessed on 14 August 2024).
55. Yao, B.; Zhang, L.; Liang, S.; Zhang, C. SVMTriP: A Method to Predict Antigenic Epitopes Using Support Vector Machine to Integrate Tri-Peptide Similarity and Propensity. *PLoS ONE* **2012**, *7*, e45152. [CrossRef] [PubMed]
56. Scheurer, S.; Son, D.Y.; Boehm, M.; Karamloo, F.; Franke, S.; Hoffmann, A.; Hausteiner, D.; Vieths, S. Cross-reactivity and epitope analysis of Pru a 1, the major cherry allergen. *Mol. Immunol.* **1999**, *36*, 155–167. [CrossRef]
57. Neudecker, P.; Lehmann, K.; Nerkamp, J.; Haase, T.; Wangorsch, A.; Fötisch, K.; Hoffmann, S.; Rösch, P.; Vieths, S.; Scheurer, S. Mutational epitope analysis of Pru av 1 and Api g 1, the major allergens of cherry (*Prunus avium*) and celery (*Apium graveolens*): Correlating IgE reactivity with three-dimensional structure. *Biochem. J.* **2003**, *376*, 97–107. [CrossRef] [PubMed]
58. Zhang, B.; Xu, S.; Liu, M.; Wei, Y.; Wang, Q.; Shen, W.; Lei, C.Q.; Zhu, Q. The nucleoprotein of influenza A virus inhibits the innate immune response by inducing mitophagy. *Autophagy* **2023**, *19*, 1916. [CrossRef]
59. Jiao, C.; Wang, B.; Chen, P.; Jiang, Y.; Liu, J. Analysis of the conserved protective epitopes of hemagglutinin on influenza A viruses. *Front. Immunol.* **2023**, *14*, 1086297. [CrossRef]
60. Corti, D.; Lanzavecchia, A. Broadly neutralizing antiviral antibodies. *Annu. Rev. Immunol.* **2013**, *31*, 705–742. [CrossRef] [PubMed]
61. Carty, M.; Bowie, A.G. Recent insights into the role of Toll-like receptors in viral infection. *Clin. Exp. Immunol.* **2010**, *161*, 397–406. [CrossRef]
62. Lester, S.N.; Li, K. Toll-like receptors in antiviral innate immunity. *J. Mol. Biol.* **2014**, *426*, 1246–1264. [CrossRef]
63. Elgert, K.D. *Immunology*; Wiley-Blackwell: Hoboken, NJ, USA, 2009.
64. Krammer, F. SARS-CoV-2 vaccines in development. *Nature* **2020**, *586*, 516–527. [CrossRef]
65. Kyriakidis, N.C.; López-Cortés, A.; González, E.V.; Grimaldos, A.B.; Prado, E.O. SARS-CoV-2 vaccines strategies: A comprehensive review of phase 3 candidates. *NPJ Vaccines* **2021**, *6*, 28. [CrossRef]
66. Martínez-Flores, D.; Zepeda-Cervantes, J.; Cruz-Reséndiz, A.; Aguirre-Sampieri, S.; Sampieri, A.; Vaca, L. SARS-CoV-2 Vaccines Based on the Spike Glycoprotein and Implications of New Viral Variants. *Front. Immunol.* **2021**, *12*, 701501. [CrossRef] [PubMed]
67. Samrat, S.K.; Tharappel, A.M.; Li, Z.; Li, H. Prospect of SARS-CoV-2 spike protein: Potential role in vaccine and therapeutic development. *Virus Res.* **2020**, *288*, 198141. [CrossRef]



68. Sulbaran, G.; Maisonnasse, P.; Amen, A.; Effantin, G.; Guilligay, D.; Dereuddre-Bosquet, N.; Burger, J.A.; Poniman, M.; Grobben, M.; Buisson, M.; et al. Immunization with synthetic SARS-CoV-2 S glycoprotein virus-like particles protects macaques from infection. *Cell Rep. Med.* **2022**, *3*, 100528. [[CrossRef](#)]
69. Wrapp, D.; Wang, N.; Corbett, K.S.; Goldsmith, J.A.; Hsieh, C.-L.; Abiona, O.; Graham, B.S.; McLellan, J.S. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **2020**, *367*, 1260–1263. [[CrossRef](#)]
70. Sternberg, A.; Naujokat, C. Structural features of coronavirus SARS-CoV-2 spike protein: Targets for vaccination. *Life Sci.* **2020**, *257*, 118056. [[CrossRef](#)]
71. Cao, Y.; Yisimayi, A.; Jian, F.; Song, W.; Xiao, T.; Wang, L.; Du, S.; Wang, J.; Li, Q.; Chen, X.; et al. BA.2.12.1, BA.4 and BA.5 escape antibodies elicited by Omicron infection. *Nature* **2022**, *608*, 593–602. [[CrossRef](#)]
72. Greaney, A.J.; Loes, A.N.; Crawford, K.H.D.; Starr, T.N.; Malone, K.D.; Chu, H.Y.; Bloom, J.D. Comprehensive mapping of mutations in the SARS-CoV-2 receptor-binding domain that affect recognition by polyclonal human plasma antibodies. *Cell Host Microbe* **2021**, *29*, 463–476.e6. [[CrossRef](#)] [[PubMed](#)]
73. Greaney, A.J.; Starr, T.N.; Barnes, C.O.; Weisblum, Y.; Schmidt, F.; Caskey, M.; Gaebler, C.; Cho, A.; Agudelo, M.; Finkin, S.; et al. Mutational Escape from the Polyclonal Antibody Response to SARS-CoV-2 Infection Is Largely Shaped by a Single Class of Antibodies. 2021. Available online: <https://www.biorxiv.org/content/10.1101/2021.03.17.435863v1> (accessed on 14 August 2024).
74. Weisblum, Y.; Schmidt, F.; Zhang, F.; DaSilva, J.; Poston, D.; Lorenzi, J.C.C.; Muecksch, F.; Rutkowska, M.; Hoffmann, H.-H.; Michailidis, E.; et al. Escape from neutralizing antibodies by SARS-CoV-2 spike protein variants. *eLife* **2020**, *9*, e61312. [[CrossRef](#)] [[PubMed](#)]
75. Röltgen, K.; Nielsen, S.C.A.; Silva, O.; Younes, S.F.; Zaslavsky, M.; Costales, C.; Yang, F.; Wirz, O.F.; Solis, D.; Hoh, R.A.; et al. Immune imprinting, breadth of variant recognition, and germinal center response in human SARS-CoV-2 infection and vaccination. *Cell* **2022**, *185*, 1025–1040.e14. [[CrossRef](#)]
76. Wheatley, A.K.; Fox, A.; Tan, H.-X.; Juno, J.A.; Davenport, M.P.; Subbarao, K.; Kent, S.J. Immune imprinting and SARS-CoV-2 vaccine design. *Trends Immunol.* **2021**, *42*, 956–959. [[CrossRef](#)]
77. Li, Y.; Ma, M.; Lei, Q.; Wang, F.; Hong, W.; Lai, D.; Hou, H.; Xu, Z.; 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](#)] [[PubMed](#)]
78. Zhang, B.; Hu, Y.; Chen, L.; Yau, T.; Tong, Y.; Hu, J.; Cai, J.; Chan, K.-H.; Dou, Y.; Deng, J.; et al. Mining of epitopes on spike protein of SARS-CoV-2 from COVID-19 patients. *Cell Res.* **2020**, *30*, 702–704. [[CrossRef](#)] [[PubMed](#)]
79. Zhang, B.-Z.; Chu, H.; Han, S.; Shuai, H.; Deng, J.; Hu, Y.; Gong, H.; Lee, A.C.-Y.; Zou, Z.; Yau, T.; et al. SARS-CoV-2 infects human neural progenitor cells and brain organoids. *Cell Res.* **2020**, *30*, 928–931. [[CrossRef](#)] [[PubMed](#)]
80. Gilbert, C.; Tengs, T. No species-level losses of s2m suggests critical role in replication of SARS-related coronaviruses. *Sci. Rep.* **2021**, *11*, 16145. [[CrossRef](#)]
81. Corman, V.M.; Muth, D.; Niemeyer, D.; Drosten, C. Hosts and Sources of Endemic Human Coronaviruses. *Adv. Virus Res.* **2018**, *100*, 163–188. [[CrossRef](#)]
82. Flerlage, T.; Boyd, D.F.; Meliopoulos, V.; Thomas, P.G.; Schultz-Cherry, S. Influenza virus and SARS-CoV-2: Pathogenesis and host responses in the respiratory tract. *Nat. Rev. Microbiol.* **2021**, *19*, 425–441. [[CrossRef](#)]
83. Traherne, J.A. Human MHC architecture and evolution: Implications for disease association studies. *Int. J. Immunogenet.* **2008**, *35*, 179–192. [[CrossRef](#)]
84. Gartland, A.J.; Li, S.; McNevin, J.; Tomaras, G.D.; Gottardo, R.; Janes, H.; Fong, Y.; Morris, D.; Geraghty, D.E.; Kijak, G.H.; et al. Analysis of HLA A\*02 Association with Vaccine Efficacy in the RV144 HIV-1 Vaccine Trial. *J. Virol.* **2014**, *88*, 8242–8255. [[CrossRef](#)]
85. Liu, Y.; Guo, T.; Yu, Q.; Zhang, H.; Du, J.; Zhang, Y.; Xia, S.; Yang, H.; Li, Q. Association of human leukocyte antigen alleles and supertypes with immunogenicity of oral rotavirus vaccine given to infants in China. *Medicine* **2018**, *97*, e12706. [[CrossRef](#)]
86. Mentzer, A.J.; O’Connor, D.; Bibi, S.; Chelysheva, I.; Clutterbuck, E.A.; Demissie, T.; Dinesh, T.; Edwards, N.J.; Felle, S.; Feng, S.; et al. Human leukocyte antigen alleles associate with COVID-19 vaccine immunogenicity and risk of breakthrough infection. *Nat. Med.* **2023**, *29*, 147–157. [[CrossRef](#)]
87. Milich, D.R.; Leroux-Roels, G.G. Immunogenetics of the response to HBsAg vaccination. *Autoimmun. Rev.* **2003**, *2*, 248–257. [[CrossRef](#)] [[PubMed](#)]
88. Nielsen, C.M.; Vekemans, J.; Lievens, M.; Kester, K.E.; Regules, J.A.; Ockenhouse, C.F. RTS,S malaria vaccine efficacy and immunogenicity during *Plasmodium falciparum* challenge is associated with HLA genotype. *Vaccine* **2018**, *36*, 1637–1642. [[CrossRef](#)] [[PubMed](#)]
89. Nishida, N.; Sugiyama, M.; Sawai, H.; Nishina, S.; Sakai, A.; Ohashi, J.; Khor, S.; Kakisaka, K.; Tsuchiura, T.; Hino, K.; et al. Key HLA-DRB1-DQB1 haplotypes and role of the BTNL2 gene for response to a hepatitis B vaccine. *Hepatology* **2018**, *68*, 848–858. [[CrossRef](#)] [[PubMed](#)]
90. O’connor, D.; Png, E.; Khor, C.C.; Snape, M.D.; Hill, A.V.; van der Klis, F.; Hoggart, C.; Levin, M.; Hibberd, M.L.; Pollard, A.J. Common Genetic Variations Associated with the Persistence of Immunity following Childhood Immunization. *Cell Rep.* **2019**, *27*, 3241–3253.e4. [[CrossRef](#)]
91. Posteraro, B.; Pastorino, R.; Di Giannantonio, P.; Ianuale, C.; Amore, R.; Ricciardi, W.; Boccia, S. The link between genetic variation and variability in vaccine responses: Systematic review and meta-analyses. *Vaccine* **2014**, *32*, 1661–1669. [[CrossRef](#)]

92. Ovsyannikova, I.G.; Haralambieva, I.H.; Vierkant, R.A.; O'Byrne, M.M.; Jacobson, R.M.; Poland, G.A. The Association of CD46, SLAM and CD209 Cellular Receptor Gene SNPs with Variations in Measles Vaccine-Induced Immune Responses: A Replication Study and Examination of Novel Polymorphisms. *Hum. Hered.* **2011**, *72*, 206. [[CrossRef](#)]
93. Arrieta-Bolaños, E.; Hernández-Zaragoza, D.I.; Barquera, R. An HLA map of the world: A comparison of HLA frequencies in 200 worldwide populations reveals diverse patterns for class I and class II. *Front. Genet.* **2023**, *14*, 866407. [[CrossRef](#)]
94. Potocnakova, L.; Bhide, M.; Pulzova, L.B. An Introduction to B-Cell Epitope Mapping and In Silico Epitope Prediction. *J. Immunol. Res.* **2016**, *2016*, 6760830. [[CrossRef](#)]
95. Aboudounya, M.M.; Heads, R.J. COVID-19 and Toll-Like Receptor 4 (TLR4): SARS-CoV-2 May Bind and Activate TLR4 to Increase ACE2 Expression, Facilitating Entry and Causing Hyperinflammation. *Mediat. Inflamm.* **2021**, *2021*, 8874339. [[CrossRef](#)]
96. Khan, S.; Shafiei, M.S.; Longoria, C.; Schoggins, J.W.; Savani, R.C.; Zaki, H. SARS-CoV-2 spike protein induces inflammation via TLR2-dependent activation of the NF- $\kappa$ B pathway. *eLife* **2021**, *10*, e68563. [[CrossRef](#)]
97. Sahanic, S.; Hilbe, R.; Dünser, C.; Tymoszyk, P.; Löffler-Ragg, J.; Rieder, D.; Trajanoski, Z.; Krogsdam, A.; Demetz, E.; Yurchenko, M.; et al. SARS-CoV-2 activates the TLR4/MyD88 pathway in human macrophages: A possible correlation with strong pro-inflammatory responses in severe COVID-19. *Heliyon* **2023**, *9*, e21893. [[CrossRef](#)] [[PubMed](#)]
98. Fontes-Dantas, F.L.; Fernandes, G.G.; Gutman, E.G.; De Lima, E.V.; Antonio, L.S.; Hammerle, M.B.; Mota-Araujo, H.P.; Colodeti, L.C.; Araújo, S.M.B.; Froz, G.M.; et al. SARS-CoV-2 Spike protein induces TLR4-mediated long-term cognitive dysfunction recapitulating post-COVID-19 syndrome in mice. *Cell Rep.* **2023**, *42*, 112189. [[CrossRef](#)] [[PubMed](#)]
99. Tyrkalska, S.D.; Martínez-López, A.; Pedoto, A.; Candel, S.; Cayuela, M.L.; Mulero, V. The Spike protein of SARS-CoV-2 signals via Tlr2 in zebrafish. *Dev. Comp. Immunol.* **2023**, *140*, 104626. [[CrossRef](#)] [[PubMed](#)]
100. Vaure, C.; Liu, Y. A comparative review of toll-like receptor 4 expression and functionality in different animal species. *Front. Immunol.* **2014**, *5*, 316. [[CrossRef](#)] [[PubMed](#)]
101. Eicher, S.D.; McMunn, K.A.; Hammon, H.M.; Donkin, S.S. Toll-like receptors 2 and 4, and acute phase cytokine gene expression in dexamethasone and growth hormone treated dairy calves. *Vet. Immunol. Immunopathol.* **2004**, *98*, 115–125. [[CrossRef](#)]
102. Su, S.B.; Tao, L.; Deng, Z.P.; Chen, W.; Qin, S.Y.; Jiang, H.X. TLR10: Insights, controversies and potential utility as a therapeutic target. *Scand. J. Immunol.* **2020**, *93*, e12988. [[CrossRef](#)]
103. Wang, Y.C.; Zhou, Y.; Fang, H.; Lin, S.; Wang, P.F.; Xiong, R.P.; Chen, J.; Xiong, X.Y.; Lv, F.L.; Liang, Q.L.; et al. Toll-like receptor 2/4 heterodimer mediates inflammatory injury in intracerebral hemorrhage. *Ann. Neurol.* **2014**, *75*, 876–889. [[CrossRef](#)]
104. Ozinsky, A.; Underhill, D.M.; Fontenot, J.D.; Hajjar, A.M.; Smith, K.D.; Wilson, C.B.; Schroeder, L.; Aderem, A. The repertoire for pattern recognition of pathogens by the innate immune system is defined by cooperation between Toll-like receptors. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 13766. [[CrossRef](#)]
105. Colleselli, K.; Stierschneider, A.; Wiesner, C. An Update on Toll-like Receptor 2, Its Function and Dimerization in Pro- and Anti-Inflammatory Processes. *Int. J. Mol. Sci.* **2023**, *24*, 12464. [[CrossRef](#)]
106. Kurt-Jones, E.A.; Popova, L.; Kwinn, L.; Haynes, L.M.; Jones, L.P.; Tripp, R.A.; Walsh, E.E.; Freeman, M.W.; Golenbock, D.T.; Anderson, L.J.; et al. Pattern recognition receptors TLR4 and CD14 mediate response to respiratory syncytial virus. *Nat. Immunol.* **2000**, *1*, 398–401. [[CrossRef](#)]
107. Ge, Y.; Mansell, A.; Ussher, J.E.; Brooks, A.E.S.; Manning, K.; Wang, C.J.H.; Taylor, J.A. Rotavirus NSP4 Triggers Secretion of Proinflammatory Cytokines from Macrophages via Toll-Like Receptor 2. *J. Virol.* **2013**, *87*, 11160. [[CrossRef](#)] [[PubMed](#)]
108. Bieback, K.; Lien, E.; Klagge, I.M.; Avota, E.; Schneider-Schaulies, J.; Duprex, W.P.; Wagner, H.; Kirschning, C.J.; ter Meulen, V.; Schneider-Schaulies, S. Hemagglutinin Protein of Wild-Type Measles Virus Activates Toll-Like Receptor 2 Signaling. *J. Virol.* **2002**, *76*, 8729. [[CrossRef](#)] [[PubMed](#)]
109. Mogensen, T.H.; Paludan, S.R. Reading the viral signature by Toll-like receptors and other pattern recognition receptors. *J. Mol. Med.* **2005**, *83*, 180–192. [[CrossRef](#)] [[PubMed](#)]
110. Lin, S.C.; Lo, Y.C.; Wu, H. Helical assembly in the MyD88:IRAK4:IRAK2 complex in TLR/IL-1R signaling. *Nature* **2010**, *465*, 885. [[CrossRef](#)]
111. Kawasaki, T.; Kawai, T. Toll-Like Receptor Signaling Pathways. *Front. Immunol.* **2014**, *5*, 461. [[CrossRef](#)]
112. Kawai, T.; Akira, S. The role of pattern-recognition receptors in innate immunity: Update on Toll-like receptors. *Nat. Immunol.* **2010**, *11*, 373–384. [[CrossRef](#)]
113. Alberts, B.; Johnson, A.; Lewis, J.; Raff, M.; Roberts, K.; Walter, P. *T Cells and MHC Proteins*; Garland Science: New York, NY, USA, 2002.
114. Boyman, O.; Sprent, J. The role of interleukin-2 during homeostasis and activation of the immune system. *Nat. Rev. Immunol.* **2012**, *12*, 180–190. [[CrossRef](#)]
115. Rokade, S.; Damani, A.M.; Oft, M.; Emmerich, J. IL-2 based cancer immunotherapies: An evolving paradigm. *Front. Immunol.* **2024**, *15*, 1433989. [[CrossRef](#)]
116. Matsushita, H.; Hosoi, A.; Ueha, S.; Abe, J.; Fujieda, N.; Tomura, M.; Maekawa, R.; Matsushima, K.; Ohara, O.; Kakimi, K. Cytotoxic T lymphocytes block tumor growth both by lytic activity and IFN $\gamma$ -Dependent Cell-cycle arrest. *Cancer Immunol. Res.* **2015**, *3*, 26–36. [[CrossRef](#)]
117. Jorgovanovic, D.; Song, M.; Wang, L.; Zhang, Y. Roles of IFN- $\gamma$  in tumor progression and regression: A review. *Biomark. Res.* **2020**, *8*, 49. [[CrossRef](#)]

118. Jawad, B.; Adhikari, P.; Podgornik, R.; Ching, W.Y. Key Interacting Residues between RBD of SARS-CoV-2 and ACE2 Receptor: Combination of Molecular Dynamics Simulation and Density Functional Calculation. *J. Chem. Inf. Model.* **2021**, *61*, 4425–4441. [[CrossRef](#)] [[PubMed](#)]
119. Borkotoky, S.; Dey, D.; Hazarika, Z. Interactions of angiotensin-converting enzyme-2 (ACE2) and SARS-CoV-2 spike receptor-binding domain (RBD): A structural perspective. *Mol. Biol. Rep.* **2023**, *50*, 2713. [[CrossRef](#)]
120. Yi, C.; Sun, X.; Ye, J.; Ding, L.; Liu, M.; Yang, Z.; Lu, X.; Zhang, Y.; Ma, L.; Gu, W.; et al. Key residues of the receptor binding motif in the spike protein of SARS-CoV-2 that interact with ACE2 and neutralizing antibodies. *Cell. Mol. Immunol.* **2020**, *17*, 621–630. [[CrossRef](#)]
121. Sun, C.; Kang, Y.F.; Liu, Y.T.; Kong, X.W.; Xu, H.Q.; Xiong, D.; Xie, C.; Liu, Y.H.; Peng, S.; Feng, G.K.; et al. Parallel profiling of antigenicity alteration and immune escape of SARS-CoV-2 Omicron and other variants. *Signal Transduct. Target. Ther.* **2022**, *7*, 42. [[CrossRef](#)]
122. Verkhivker, G.; Alshahrani, M.; Gupta, G. Balancing Functional Tradeoffs between Protein Stability and ACE2 Binding in the SARS-CoV-2 Omicron BA.2, BA.2.75 and XBB Lineages: Dynamics-Based Network Models Reveal Epistatic Effects Modulating Compensatory Dynamic and Energetic Changes. *Viruses* **2023**, *15*, 1143. [[CrossRef](#)]
123. Kumar, S.; Thambiraja, T.S.; Karuppanan, K.; Subramaniam, G. Omicron and Delta variant of SARS-CoV-2: A comparative computational study of spike protein. *J. Med. Virol.* **2022**, *94*, 1641–1649. [[CrossRef](#)]
124. Wu, X.; Li, W.; Rong, H.; Pan, J.; Zhang, X.; Hu, Q.; Shi, Z.-L.; Zhang, X.-E.; Cui, Z. A Nanoparticle Vaccine Displaying Conserved Epitopes of the Preexisting Neutralizing Antibody Confers Broad Protection against SARS-CoV-2 Variants. *ACS Nano* **2024**, *18*, 17749–17763. [[CrossRef](#)] [[PubMed](#)]
125. He, L.; Lin, X.; Wang, Y.; Abraham, C.; Sou, C.; Ngo, T.; Zhang, Y.; Wilson, I.A.; Zhu, J. Single-component, self-assembling, protein nanoparticles presenting the receptor binding domain and stabilized spike as SARS-CoV-2 vaccine candidates. *Sci. Adv.* **2021**, *7*, eabf1591. [[CrossRef](#)] [[PubMed](#)]
126. Ng, A.K.-L.; Zhang, H.; Tan, K.; Li, Z.; Liu, J.; Chan, P.K.-S.; Li, S.-M.; Chan, W.-Y.; Au, S.W.-N.; Joachimiak, A.; et al. Structure of the influenza virus A H5N1 nucleoprotein: Implications for RNA binding, oligomerization, and vaccine design. *FASEB J.* **2008**, *22*, 3638. [[CrossRef](#)]
127. Wu, F.; Huang, J.-H.; Yuan, X.-Y.; Huang, W.-S.; Chen, Y.-H. Characterization of immunity induced by M2e of influenza virus. *Vaccine* **2007**, *25*, 8868–8873. [[CrossRef](#)]
128. Gao, X.; Wang, W.; Li, Y.; Zhang, S.; Duan, Y.; Xing, L.; Zhao, Z.; Zhang, P.; Li, Z.; Li, R.; et al. Enhanced Influenza VLP vaccines comprising matrix-2 ectodomain and nucleoprotein epitopes protects mice from lethal challenge. *Antiviral Res.* **2013**, *98*, 4–11. [[CrossRef](#)] [[PubMed](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.