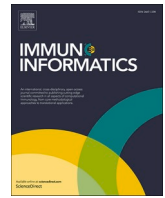




Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



SARS-CoV-2 Omicron (BA.1 and BA.2) specific novel CD8⁺ and CD4⁺ T cell epitopes targeting spike protein

Simone Parn^a, Kush Savsani^b, Sivanesan Dakshanamurthy^{c,*}

^a University of the District of Columbia, Washington, D.C, USA

^b Virginia Commonwealth University, Richmond, VA, USA

^c Lombardi Comprehensive Cancer Center, Georgetown University Medical Center, Washington DC 20057 USA

ARTICLE INFO

Keywords:

Epitope-based vaccine
ImmunoInformatics
SARS-CoV-2 Omicron
COVID-19 variants
Spike glycoprotein

ABSTRACT

The Omicron (BA.1/B.1.1.529) variant of SARS-CoV-2 harbors an alarming 37 mutations on its spike protein, reducing the efficacy of current COVID-19 vaccines. In this study, we identified CD8⁺ and CD4⁺ T cell epitopes from SARS-CoV-2 S protein mutants. To identify the highest quality CD8 and CD4 epitopes from the Omicron variant, we selected epitopes with a high binding affinity towards both MHC I and MHC II molecules. We applied other clinical checkpoint predictors, including immunogenicity, antigenicity, allergenicity, instability and toxicity. Subsequently, we found eight Omicron (BA.1/B.1.1.529) specific CD8⁺ and eleven CD4⁺ T cell epitopes with a world population coverage of 76.16% and 97.46%, respectively. Additionally, we identified common epitopes across Omicron BA.1 and BA.2 lineages that target mutations critical to SARS-CoV-2 virulence. Further, we identified common epitopes across B.1.1.529 and other circulating SARS-CoV-2 variants, such as B.1.617.2 (Delta). We predicted CD8 epitopes' binding affinity to murine MHC alleles to test the vaccine candidates in preclinical models. The CD8 epitopes were further validated using our previously developed software tool PCOptim. We then modeled the three-dimensional structures of our top CD8 epitopes to investigate the binding interaction between peptide-MHC and peptide-MHC-TCR complexes. Notably, our identified epitopes are targeting the mutations on the RNA-binding domain and the fusion sites of S protein. This could potentially eliminate viral infections and form long-term immune responses compared to relatively short-lived mRNA vaccines and maximize the efficacy of vaccine candidates against the current pandemic and potential future variants.

Introduction

The novel Omicron (BA.1/B.1.1.529) variant of SARS-CoV-2, first identified in South Africa in 2021 [1], was quickly declared as a variant of concern (VOC) as it had the potential to infect many populations across the world. Although several COVID-19 vaccines and antiviral drugs are currently available, the new variant has challenged public health at an unprecedented scale. Characterized by a heavy mutation load, Omicron has raised alarming immune escape and potentially increased transmissibility compared to the Delta variant [2]. B.1.1.529 harbors an alarming number of spike protein substitutions, with a total of 15 located on the RNA-binding domain (RBD) of the S protein alone, meaning the new VOC may confer enhanced binding to the human Angiotensin-Converting Enzyme 2 (ACE2) receptor [2,3]. Moreover, the recent emergence and spread of Omicron sub-lineage BA.2 may pose

another risk to public safety. Thus, there is an urgent need for effective booster vaccines against SARS-CoV-2 Omicron variants to reduce the spread of this highly infectious agent.

SARS-CoV-2 spike (S) glycoprotein is a first-generation COVID-19 target due to its structure and function, which play an important role in viral infection and pathogenesis [4]. Mature S protein consists of two subunits: S1, which includes the RNA-binding domain responsible for the interaction with ACE2, and S2, which facilitates membrane fusion and viral entry [5,6]. The S1 subunit contains the N-terminal domain (NTD) and the RBD, which are the major targets of polyclonal and monoclonal neutralizing antibodies [7]. In addition, SARS-CoV-2 hosts a novel protease cleavage site at S1/S2 predicted to increase the virulence of the virus by enabling the cleavage of fusion peptides and therefore promoting entry into lung cells [4]. The continued spread of COVID-19 in immunocompromised individuals allows the virus to mutate while

* Corresponding author.

E-mail address: sd233@georgetown.edu (S. Dakshanamurthy).

<https://doi.org/10.1016/j.immuno.2022.100020>

generating new SARS-CoV-2 variants. The emergence of B.1.1.529 S protein point mutations and deletions, such as mutations at residues 142–145, 417, 484, and 501, may be correlated with increased viral transmission and resistance to NTD or RBD antibodies, therefore creating speculation whether the current vaccines provide efficient protection against the virus [8,9].

The need for a quick and effective vaccine to combat existing and emerging infectious diseases has implemented the use of *in silico* prediction methods, such as immunoinformatics [10]. Specifically, immunoinformatics provides a low-cost and rapid approach for epitope-based vaccine design which has many advantages over other conventional types of vaccines. Epitope-based vaccines offer more specific immune responses, promote long-lasting immunity and reduce undesired side effects [11]. Memory T cells specific for SARS-CoV-2 epitopes can persist for 11 years after infection [12] compared to relatively short-lived antibody responses [13]. Moreover, T cell responses in transgenic mice have been correlated with similar T cell responses in vaccinated and infected humans [14,15], indicating the potential of epitope-based vaccines to undergo clinical testing at an accelerated rate.

In this study, we obtained immunogenic CD8+ and interferon-gamma (IFN γ) inducing CD4+ T cell epitopes from spike protein of SARS-CoV-2 for a multi-epitope vaccine to protect against the circulating COVID-19 variants and Omicron variant, specifically. Immune Epitope Database (IEDB) [16] was utilized to predict CD8+ and CD4+ T cell epitopes with high binding affinity to MHC class I and II alleles, respectively. Beyond MHC affinity, epitopes were further refined by predicting immunogenicity, population coverage, antigenicity, allergenicity, toxicity, IFN γ secretion, half-life and GRAVY (grand average of hydropathicity), and other amino acid physicochemical properties to fully maximize the quality and efficiency of the designed vaccine. After predicting CD4+ and CD8+ T cell epitopes, we attempted to identify top immunogenic CD8 peptides (9–10 aa long) within IFN γ inducing CD4 epitopes (15 aa long) that could be used to provoke a strong and long-lasting immune response against the Omicron variant. We validated the predicted epitopes using our previously developed program called PCOptim [17]. We then predicted the 3D structures of the final CD8 T cell epitopes and docked them with their respective HLA alleles to visualize the peptide-MHC interaction. Since T-cell receptor (TCR) in complex with pMHC structure is thought to play an important role in the immune response against evading pathogens, we also visualized the pMHC-TCR interaction to gain a better insight into the vaccine mechanism. To safely test the peptide-based vaccine constructs for their efficacy and immunogenicity in pre-clinical trials, we also predicted murine MHC binding affinity to our CD8 peptides.

Materials and methods

Retrieval of SARS-CoV-2 sequence

We retrieved the reference S protein sequence of SARS-CoV-2 Wuhan isolates from the NCBI database using accession number YP_009724390.1. The Omicron BA.1 and BA.2 specific S protein sequences were generated using lineage-defining S protein mutations [18, 19].

CD8+ t cell epitope prediction and immunogenicity modeling

CD8+ T cell epitopes were predicted using NetMHCpan EL 4.1 (IEDB Recommended) [20–24]. We predicted CD8+ T cell epitopes for the frequently occurring MHC-I-binding alleles and the HLA allele reference set. The amino acid length of peptide 9.0 and 10.0 were selected as parameters for identifying MHC I alleles. Subsequently, we selected the top 300 unique epitopes based on rank and predicted their immunogenicity using the IEDB Immunogenicity tool on <http://tools.iedb.org/immunogenicity/> [25]. Epitopes with an immunogenicity score >0 were selected for further analysis.

CD4+ t cell epitope prediction

We used the IEDB Analysis Resource v2.22 to predict CD4+ T cell epitopes [26,27]. Full HLA reference set and default epitope length were selected as parameters for predicting MHC II binders. We chose the top 300 unique epitopes based on rank for further analysis.

IFN γ inducing CD4 peptide prediction

Due to the lack of reliable immunogenicity predictor on IEDB.org, the top 300 unique CD4 peptides were refined based on IFN γ inducer properties using the IFNEpitope server (<https://webs.iitd.edu.in/raghava/ifnepitope/index.php>) [28]. IFNEpitope is designed to make predictions based on peptide length, positional conservation of residues and amino acid composition. Using a hybrid approach of the motif and SVM-based predictions, epitopes returned either positive or negative induction for IFN γ release. Only epitopes that were “positive” for IFN γ release were selected for further analysis.

Antigenicity prediction

VaxiJen v2.0 [29] was used to predict antigenic CD8 and CD4 epitopes (<http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html>). VaxiJen predicts epitopes based on the physicochemical properties of the protein sequences. A viral model was used, for predictions and epitopes that returned a score >0.4 were considered antigenic.

Allergenicity prediction

AllerTop v2.0 [30] was used to predict allergenicity for both CD8 and CD4 epitopes (<http://www.pharmfac.net/allertop>). The server provides an alignment-free method for *in silico* prediction of allergens and non-allergens based on the physicochemical properties of epitopes. Only “probable non-allergen” peptides were selected for our analysis.

Toxicity prediction

The toxicity of CD8 and CD4 peptides was determined using the ToxinPred server on <https://webs.iitd.edu.in/raghava/toxinpred/> [31]. The method was developed based on machine learning and a quantitative matrix using several properties of peptides. The tool returns both toxic and nontoxic peptides. For our analysis, only “non-toxin” epitopes were selected.

Amino acid physicochemical properties

Evaluation of various physicochemical properties is essential for determining the safety and efficacy of the candidate vaccine. Using ProtParam [32] tool on ExPASy (<https://web.expasy.org/protparam/>), the chemical and physical properties of the top CD8 and CD4 peptides were assessed. Instability index, aliphatic index, GRAVY and half-life were among the parameters used in the study. The instability index was calculated to determine whether an epitope was stable or unstable *in vivo* where an instability index <40 was selected as the threshold for stable epitopes. Half-life in mammalian reticulocytes *in vitro* was predicted where half-life >1 hour was chosen as a threshold value.

Worldwide human population coverage analysis

We calculated the population coverage of predicted epitopes with known MHC restriction using the IEDB population coverage tool on <http://tools.iedb.org/population/> [33]. The percentage of the world population predicted to present our epitopes on their MHC molecules was computed by the tool. The population coverage analysis was performed separately for MHC I and II binders. The analysis covered East Asia, Northeast Asia, South Asia, Southeast Asia, Southwest Asia, Europe, East

Africa, West Africa, Central Africa, North Africa, South Africa, West Indies, North America, Central America, South America and Oceania. Only HLA alleles from the top 300 predicted epitopes ordered according to “rank” were considered in the analysis.

Murine mhc restriction prediction

We utilized SYFPEITHI on <http://www.syfpeithi.de> [34] to predict the CD8 peptide’s affinity to murine MHC class I (H2) molecules. The prediction tool is based on published T cell epitopes and MHC ligands. It takes into consideration the amino acids in the anchor and auxiliary anchor positions, as well as other frequent amino acids. Two common mouse strains to study COVID-19 vaccines, C57BL/6 and BALB/CJ, were used in mouse MHC affinity predictions. Since scores calculated by the SYFPEITHI server vary, we identified peptides that have been experimentally validated to bind to C57BL/6 and BALB/CJ supertypes. Subsequently, we compared the scores of reference peptides: RTFSFQLI [35], MYIFPVHWQF [36] and RPQASGVYM [37] to our predicted CD8 epitopes.

Validation using our population coverage optimization software

We earlier developed a program called PCOptim to validate the predicted epitopes. The program can be used to generate an optimized dataset of epitopes with maximum population coverage [17]. The tool accepts the input of several epitope-HLA allele pairs and generates an optimized list of epitopes that provide the maximum possible population coverage. PCOptim produces the optimized list of epitopes by conserving all unique HLA alleles, thereby ensuring that population coverage remains maximized. However, the program does not consider several clinical checkpoint variables: antigenicity, allergenicity, or toxicity. The epitopes that pass all clinical checkpoint parameters will be more limited and likely have lower population coverage than the optimized dataset. Therefore, the optimized set of epitopes can be used as a point of comparison to identify geographic regions that may receive lower coverage.

Three-dimensional (3D) structure prediction

Three-dimensional (3D) structure prediction is useful in illustrating the binding of the peptide to a respective MHC molecule. To visualize the peptide-MHC complex, we first predicted the tertiary structure of the top CD8 epitopes using PEPstrMOD [38]. The 3D structure of the HLA allele that the S protein epitope binds to was downloaded from Protein Data Bank on [RCSB.org](https://www.rcsb.org). The native peptides from the RCSB structures were replaced with our predicted epitopes using PyMol. We then spatially docked the predicted epitope in the MHC groove using PyMOL and FlexPepDock [39]. The best model provided by the server was energy minimized using HADDOCK 2.4 [40]. We also modeled the 3D structure of the peptide-MHC-TCR complex using TCRmodel [41].

Results

Workflow of t cell predictions

A schematic presentation of the immunoinformatic methods was proposed for epitope predictions as described in our previous study (Fig. 1) [42]. We started the analysis by creating SARS-CoV-2 Omicron (BA.1 and BA.2) specific S protein sequences where lineage-defining mutations were replaced in the SARS-CoV-2 Wuhan reference sequence. We used experimentally validated prediction tools from our previous study [43] to predict CD8 and CD4 epitopes with MHC Class I and II binding affinity, respectively. We applied additional parameters to obtain the highest quality epitopes for the multi-epitope vaccine. World population coverage was thereafter calculated to assess the protection scope of selected epitopes. To test the efficacy of the epitopes in

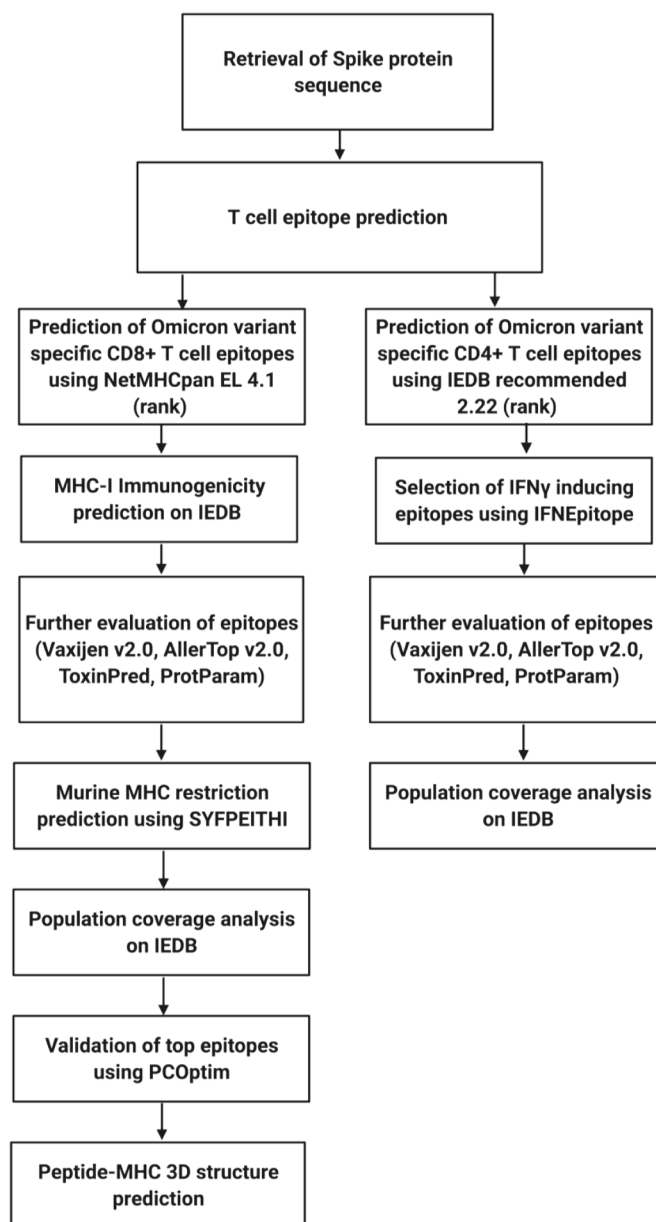


Fig. 1. Flowchart of the T cell epitope prediction mechanism. Omicron variant specific CD8+ and CD4+ T cell epitopes were predicted using freely available prediction tools on Immune Epitope Database (IEDB). MHC I binders were refined by immunogenicity and MHC II binders by IFN γ inducer capability. All CD8 and CD4 epitopes were then evaluated and selected based on their antigenicity, allergenicity, toxicity and physiochemical properties. Population coverage was computed for both CD8 and CD4 epitopes with known MHC restriction. CD8 peptides were further validated using our software PCOptim. Murine MHC binding affinity to top CD8 epitopes was predicted and the peptide-MHC structures were modeled in three-dimensional analysis. The figure was created with BioRender.com.

preclinical trials, we predicted the binding affinity of our CD8 epitopes to murine MHC alleles. To validate the CD8 epitopes, we utilized our earlier developed program called PCOptim. Eventually, we visualized peptide-MHC (pMHC) and pMHC-TCR interactions in three-dimensional models to give a better insight into T cell-mediated immune response mechanism.

SARS-CoV-2 omicron variant mutations

The Omicron variant, also known as PANGO lineage BA.1 or

B.1.1.529, harbors up to 59 mutations in the entire genome with 37 of them located on the S protein, the mediator of host cell entry (Fig. 2). Among these, Omicron hosts a novel three amino-acid insertion at position 214. The RBD region, considered the main target of neutralizing antibodies as it binds to the human ACE2 receptor, contains 15 of these mutations. Besides novel mutations in the Omicron BA.1 variant, some of the mutations are also observed in other SARS-CoV-2 variants. The Omicron sub-lineages BA.1 and BA.2 differ in some of the mutations, including the spike protein, however, the two have at least 20 common mutations (Fig. 2). B.1.617.2 variant shares three amino acid mutations with B.1.1.529: G142D, T478K and D614G, whereas B.1.351 and P.1 share mutations in residues K417N, E484K, and N501Y, which are correlated with increased binding affinity to ACE2 [44,45].

CD8+ t cell epitope prediction

We used NetMHCpan EL 4.1 based on rank and additional parameters including MHC-I immunogenicity predictor, VaxiJen 2.0, AllerTop 2.0, ToxinPred and ProtParam, to select only immunogenic, antigenic, non-allergenic, nontoxic and stable CD8+ T cell epitopes from SARS-CoV-2 Omicron (BA.1) S protein. The search yielded 25 immunogenic, antigenic, non-allergenic, nontoxic and stable epitopes. We then compared these epitopes to other SARS-CoV-2 variant CD8 peptides identified by our previous study [43]. This led us to identify 12 of 25 common epitopes among Omicron (BA.1/B.1.1.529), Alpha (B.1.1.7), Beta (B.1.351), Gamma (P.1), Delta (B.1.617.2), US variants (S protein mutations), and Cluster 5 mink variants (Supplementary Table S1) which could be used in a multi-epitope vaccine to target multiple variants simultaneously. The remaining 13 CD8 peptides appeared to match

the S protein reference sequence of SARS-CoV-2 indicating the absence of any BA.1 variant specific epitopes.

The stringent method provides an efficient way to match epitopes with 100% sequence similarity. We compared all predicted immunogenic reference epitopes to immunogenic BA.1 epitopes. The search successfully returned 25 BA.1 specific immunogenic epitopes where eight were identified as immunogenic, antigenic, non-allergenic, nontoxic but unstable in vivo (Table 2, Supplementary Table S2). We also predicted CD8 peptides specific to the stealth version of the Omicron variant of the BA.2 sub-lineage. The study yielded three immunogenic antigenic non-allergenic nontoxic and stable epitopes specific to BA.2 sub-lineage (Supplementary Table S3). Using our 8 BA.1 specific CD8 epitopes, we aimed to identify any common epitopes among BA.1 and BA.2 sub-lineages. The search yielded one CD8+ T cell epitope (KSHRRARSV) present in both Omicron lineages and located at the furin cleavage site. Finally, we checked if there were any BA.2 CD8 epitopes were also common to other SARS-CoV-2 strains. Successively, BA.2 shares identical epitopes with all other SARS-CoV-2 variants listed in Supplementary Table S1.

CD4+ t cell epitope prediction

Using the IEDB Analysis Resource v2.22 based on rank, IFNEpitope, VaxiJen 2.0, AllerTop 2.0 and ToxinPred, only IFN γ inducing antigenic non-allergenic nontoxic CD4 epitopes were selected. Our analysis yielded seven common S CD4+ T cell epitopes across Omicron (BA.1/B.1.1.529) and six other SARS-CoV-2 variants (Alpha, Beta, Delta, Gamma, US variants (S protein mutations), and Cluster 5 mink variants) reported by our other study (Table 3, Supplementary Table S4) [43]. In

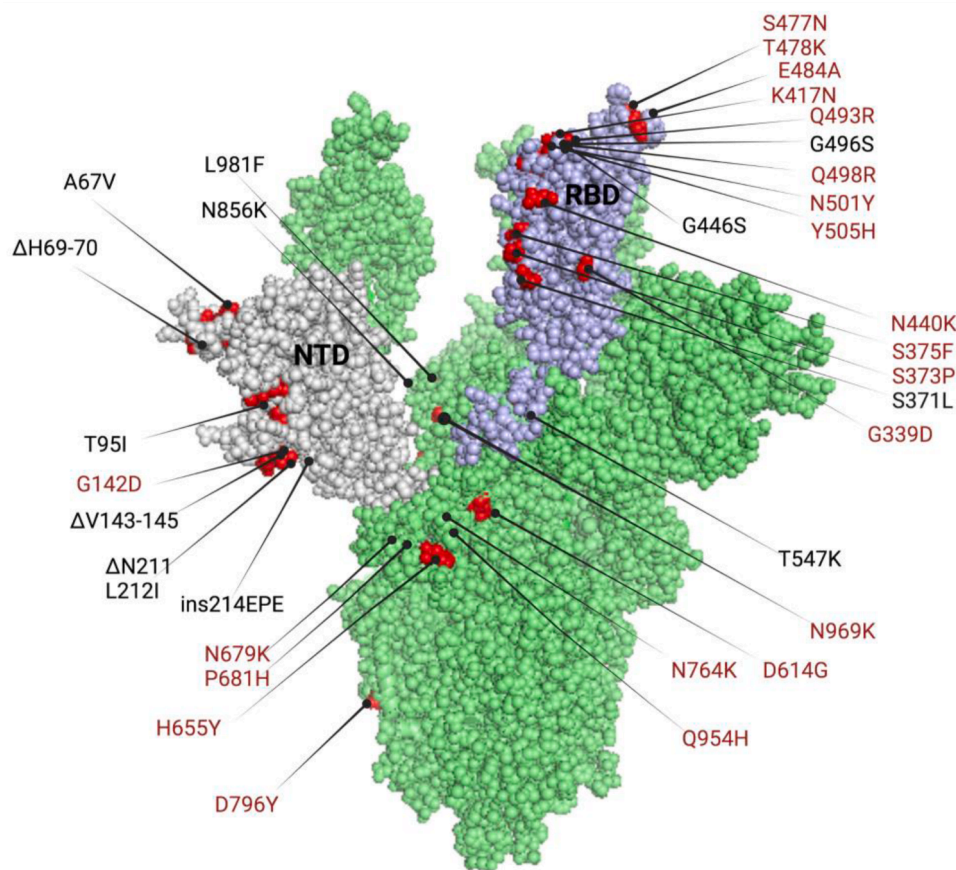


Fig. 2. 3D structure of SARS-CoV-2 S protein (PDB: 7VNE) highlighting the mutational landscape of Omicron variant (BA.1/ B.1.1.529). Omicron variant harbors 37 mutations on the S protein with most of the mutations located on the RNA-binding domain. Mutations highlighted in red are shared between BA.1 and BA.2 sub-lineages.

addition, we identified 15 Omicron BA.1 specific S CD4+ T cell epitopes (Supplementary Table S5), whereas 11 of them were predicted to be stable in vivo. We also predicted CD4+ T cell epitopes from the BA.2 variant to check any common epitopes among Omicron BA.1 specific CD4 peptides and across other SARS-CoV-2 strains. Our analysis resulted in 8 common CD4+ T cell epitopes across Omicron BA.1 and BA.2 sub-lineages (Tables 2 and 3). Moreover, BA.2 shares identical CD4 peptides with other SARS-CoV-2 variants listed in Table 3.

Identification of overlapping t cell epitopes

CD8+ T cells are important for the clearance of infectious cells and CD4+ T cells play a crucial role in promoting specific B cell antibody production. It is thought that together these T cells can elicit a robust immune response against an evading pathogen. Hence, we overlapped the CD8+ T cell epitopes on the CD4+ T cell epitopes to identify overlapping Omicron BA.1 specific epitopes. The analysis from the previous section revealed 3 S CD4+ T cell epitopes (KSHRRARSVASQSII, SVLYNLAPFFTFKCY and VLYNLAPFFTFKCYG) overlapping with our top CD8+ T cell epitopes (Table 3). All identified CD4+ T cell epitopes are antigenic, IFN γ inducing, non-allergenic, and nontoxic. Moreover, we found that CD4+ T cell (KSHRRARSVASQSII) that overlaps with a CD8+ T cell epitope (KSHRRARSV) is common to both Omicron BA.1 and BA.2 sub-lineages.

Murine mhc restriction prediction

To assess the efficacy of the proposed vaccine candidates, we predicted their binding affinity to murine H2 alleles. Since the SYFPEITHI server is mostly validated for MHC Class I epitope prediction and has low reliability (50%) for MHC Class II peptides, we only identified CD8 peptides that bind to murine H2 alleles. The comparison of reference epitopes to the identified CD8 peptides yielded potential strong, intermediate and weak binders. Our analysis revealed seven common S CD8+ T cell epitopes among SARS-CoV-2 variants (Supplementary Table S1) and four Omicron BA.1 specific S CD8+ T cell epitopes (Table 1) presented by murine MHC restriction. Moreover, we identified one BA.2 lineage specific CD8+ T cell epitope (TPINLGRDL) presented by murine MHC restriction (Supplementary Table S3). Ultimately, the predicted strong, intermediate and weak murine MHC binders can be used in further preclinical trials on mouse models with similar allelic presentation.

Population coverage analysis

The distribution of MHC alleles varies across geographical areas and ethnicities. Thus, it is highly important to consider population coverage when designing an effective vaccine. To determine the world population

coverage of the selected T cell epitopes, corresponding HLA associations were considered. Epitopes that bind several MHC (HLA) alleles are generally regarded as the best probable epitopes as they have a higher potential to show good coverage by approaching 100%. We found that the 12 common CD8 peptides among Omicron and other SARS-CoV-2 variants cover 84.0% of the world population (Supplementary Fig. 1A). In contrast, the eight Omicron BA.1 specific CD8 peptides cover 76.16% (Table 1, Supplementary Fig. 1B). World population coverage was also computed for the MHC Class II peptides. We found that the seven common CD4+ T cell epitopes across Omicron and other circulating variants, including Delta, can elicit an immune response that covers 96.65% of the world population (Supplementary Fig. 1C). The 11 top-quality Omicron BA.1 variant specific CD4 peptides were computed to cover 97.46% of the world population (Table 2, Supplementary Fig. 1D) whereas the three overlapping CD4 peptides were found to cover 92.66% of the world population (Table 3, Supplementary Fig. 1E).

Validation of top epitopes

The top 300 epitopes found from IEDB NetMHCpan EL 4.1 have a rank range of 0–0.30. We filtered through all the raw epitope data to obtain a dataset of epitopes with a rank less than 0.30 and immunogenicity greater than 0. We entered this dataset into PCOptim to obtain a list of optimized epitopes, which we can use for comparison to our eight immunogenic, antigenic, non-allergenic, nontoxic and stable epitopes. The optimized list of epitopes included eight unique epitopes and returned 98.55% population coverage (Table 4, Fig. 3). Of the eight optimized epitopes, two are present in our “Omicron BA.1 variant specific S CD8 epitope” dataset (Supplementary Table S2) (RSYSFRPTY and YNLAPFFTF) and one is present in our “Top common S CD8 epitopes across Omicron and other SARS-CoV-2 variants” dataset (Supplementary Table S1) (VVFLHVTYV). The HLA alleles present in the optimized dataset but not in the clinically relevant dataset are: HLA-B*07:02, HLA-B*08:01, HLA-A*31:01, HLA-A*33:01, and HLA-B*51:01. These alleles account for the difference in population coverage between the optimal epitope dataset and datasets obtained after passing epitopes through several clinical checkpoints.

Three-dimensional (3D) structural analysis

Protein 3D structures offer useful insights into their molecular activity and provide a wide variety of applications in bioscience. We selected two immunogenic antigenic non-allergenic and nontoxic CD8+ T cell epitopes (KSHRRARSV and YNLAPFFTF) for 3D visualization and molecular docking analysis. KSHRRARSV and YNLAPFFTF are both strong MHC I binders and presented by murine MHC restriction. The selected CD8 epitopes were docked individually with the alleles they were restricted to, which was HLA-A*30:01 (RCSB PDB: 6J1W) for

Table 1

Omicron (BA.1/B.1.1.529) specific S protein CD8+ T cell epitopes and their world population coverage. All predicted epitopes are immunogenic, antigenic, non-allergenic and nontoxic. Variant specific mutations are written in red color. We identified one common CD8+ T cell epitope (KSHRRARSV) across Omicron BA.1 and BA.2 sub-lineages. Murine model MHC restriction is also indicated for peptides that were identified as weak, intermediate or strong binders.

World Population Coverage	Omicron Specific CD8 Peptide	Targeted S Mutations	HLA Restriction	Mouse MHC Restriction	Omicron sub-lineage
76.16%	NLAPFFTFK	S371L, S373P, S375F	HLA-A*11:01, HLA-A*03:01, HLA-A*68:01	Not Available	BA.1
	YNLAPFFTF	S371L, S373P, S375F	HLA-A*23:01, HLA-A*24:02	H2-Ld, H2-Db	BA.1
	GVYFASIEK	T95I	HLA-A*11:01, HLA-A*03:01, HLA-A*30:01	Not Available	BA.1
	KSHRRARSV	N679K, P681H	HLA-A*30:01	H2-Kd	BA.1, BA.2
	APFFTFKCY	S373P, S375F	HLA-B*35:01	H2-Ld	BA.1
	VLYNLAPFF	S371L, S373P, S375F	HLA-A*32:01, HLA-A*23:01	Not Available	BA.1
	REPEDLPQGF	ins214EPE	HLA-B*44:03, HLA-B*44:02	H2-Db	BA.1
	RSYSFRPTY	Q493R, G496S, Q498R, N501Y	HLA-A*30:02, HLA-A*32:01, HLA-A*30:01, HLA-B*57:01, HLA-B*15:01, HLA-B*58:01, HLA-A*03:01, HLA-A*11:01	Not Available	BA.1

Table 2

Top Omicron (BA.1/B.1.1.529) specific S protein CD4 peptides and their population coverage. All epitopes are IFN γ inducing, antigenic, non-allergenic, nontoxic and stable. Variant specific mutations are written in red color. The world population coverage of the selected epitopes is also presented. We identified seven common CD4+ T cell epitopes across Omicron BA.1 and BA.2 sub-lineages.

World Population Coverage	Omicron Specific CD4 Peptide	Targeted S mutations	HLA Restriction	Omicron sub-lineage
97.46%	FLPFFSNVTWFHVIS	A67V, del69-70	HLA-DPA1*01:03, HLA-DPB1*04:01, HLA-DRB3*02:02, HLA-DPA1*01:03, HLA-DPB1*02:01	BA.1
	LPPFFSNVTWFHVISG	A67V, del69-70	HLA-DPA1*01:03, HLA-DPB1*04:01, HLA-DPA1*01:03, HLA-DPB1*02:01, HLA-DRB3*02:02	BA.1
	TQLKRALTGIAVEQD	N764K	HLA-DPA1*02:01, HLA-DPB1*14:01, HLA-DQA1*03:01, HLA-DQB1*03:02, HLA-DQA1*04:01, HLA-DQB1*04:02	BA.1, BA.2
	FCTQLKRALTGIAVE	N764K	HLA-DPA1*02:01, HLA-DPB1*14:01	BA.1, BA.2
	SNLLQYGSFCTQLK	N764K	HLA-DRB1*15:01	BA.1, BA.2
	LKRALTGIAVEQDKN	N764K	HLA-DPA1*02:01, HLA-DPB1*14:01, HLA-DQA1*04:01, HLA-DQB1*04:02, HLA-DQA1*03:01, HLA-DQB1*03:02	BA.1, BA.2
	GSFCTQLKRALTGIA	N764K	HLA-DRB1*11:01	BA.1, BA.2
	YGSFCTQLKRALTGI	N764K	HLA-DRB1*11:01	BA.1, BA.2
	KRALTGIAVEQDKNT	N764K	HLA-DQA1*03:01, HLA-DQB1*03:02	BA.1, BA.2
	SVLYNLAPFFTFKCY	S371L, S373P, S375F	HLA-DPA1*01:03, HLA-DPB1*04:01, HLA-DPA1*01:03, HLA-DPB1*02:01	BA.1
	VLYNLAPFFTFKCYG	S371L, S373P, S375F	HLA-DPA1*01:03, HLA-DPB1*04:01, HLA-DPA1*01:03, HLA-DPB1*02:01	BA.1

Table 3

Omicron (BA.1/B.1.1.529) specific IFN γ inducing, antigenic, non-allergenic and nontoxic CD4 peptides which consist of immunogenic non-allergenic antigenic nontoxic CD8 peptides. The world population coverage of the selected epitopes is also presented. Variant specific mutations are written in red color. We identified one overlapping CD4 and CD8 peptide (KSHRRARSVASQSII) across Omicron BA.1 and BA.2 sub-lineages.

World Population Coverage	Omicron Specific CD4 Peptide	Omicron Specific CD8 Peptide	Targeted S mutation	MHCII Restriction	Instability
92.66%	SVLYNLAPFFTFKCY	NLAPFFTFKCY YNLAPFFTF	S371L, S373P, S375F	HLA-DPA1*01:03, HLA-DPB1*04:01, HLA-DPA1*01:03, HLA-DPB1*02:01	Stable
	VLYNLAPFFTFKCYG	NLAPFFTFKCY YNLAPFFTF	S371L, S373P, S375F	HLA-DPA1*01:03, HLA-DPB1*04:01, HLA-DPA1*01:03, HLA-DPB1*02:01	Stable
	KSHRRARSVASQSII	KSHRRARSV	N679K, P681H	HLA-DPA1*02:01, HLA-DPB1*14:01	Unstable

Table 4

Optimized list of epitopes returned by PCOptim. Input data included all raw epitopes with a rank less than 0.30 and immunogenicity greater than 0. HLA alleles that are bolded are also present in the HLA restrictions of epitopes that passed all clinical checkpoint parameters. RSYSFRPTY and YNLAPFFTF are present in Omicron variant specific S CD8 epitopes and VVFLHVTYV is present in Top common epitopes across Omicron and other SARS-CoV-2 variants.

Peptide	HLA Restriction
SVYAWNRRK	HLA-A*31:01, HLA-A*33:01, HLA-A*68:01, HLA-A*03:01, HLA-A*11:01
VVFLHVTYV	HLA-A*02:03, HLA-A*02:06, HLA-A*68:02, HLA-A*02:01
AEIRASANL	HLA-B*40:01, HLA-B*44:03, HLA-B*44:02
RSYSFRPTY	HLA-A*30:02, HLA-A*32:01, HLA-A*30:01, HLA-B*57:01, HLA-B*15:01, HLA-B*58:01, HLA-A*03:01, HLA-A*11:01
YNLAPFFTF	HLA-A*23:01, HLA-A*24:02
QPTESIVRF	HLA-B*35:01, HLA-B*53:01, HLA-B*51:01, HLA-B*07:02
YLQPRTEFL	HLA-A*02:01, HLA-B*08:01, HLA-A*02:03, HLA-A*02:06, HLA-A*32:01
LTDEMIAQY	HLA-A*01:01, HLA-A*30:02, HLA-A*26:01, HLA-B*35:01

epitope KSHRRARSV (Fig. 4A) and HLA-A*24:02 (RCSB PDB: 5WWI) for epitope YNLAPFFTF (Fig. 4B). The structures were also energy minimized to reduce the overall potential energy between the epitope and MHC molecule. Furthermore, to get a better insight into the TCR interaction with the pMHC complex, we modeled the YNLAPFFTF-HLA-A*24:02 complex with the C1-28 TCR (RCSB PDB: 3VXM) structure specific to HLA-A24 (Fig. 4C).

Discussion

Although several COVID-19 vaccines are currently available, the excessive mutations observed in the spike protein of Omicron can escape

immune response, raising concerns over the efficacy of these vaccines [44,46]. Despite the overall success of mRNA vaccines which have provided breakthrough technology in protecting the community against the severity of infection, it remains urgent to expand the landscape of vaccines to combat the novel Omicron variant as well as any future SARS-CoV-2 strains [47,48]. Here, we present an epitope-based vaccine design which has many advantages over the current mRNA vaccines. Epitope-based vaccines are highly specific and able to target multiple SARS-CoV-2 antigens simultaneously. They also confer a long-lasting immune response, superior to the relatively short-lived mRNA vaccines, which have shown to decline after six months of vaccination [49]. Another important advantage of epitope-based vaccines is their absence of adverse side effects, which avoid allergenic consequences.

We carefully selected the prediction tools based on their accuracy to design an efficient and safe vaccine that would provide broad protection across many ethnicities. The IEDB prediction servers are free and widely accepted in literature [50]. However, to choose the most reliable T cell prediction tool, we used our previous study that analyzed the specificity and sensitivity of several MHC I and MHC II peptide prediction tools [43]. We used NetMHCpan EL 4.1 for CD8 epitope predictions, and the IEDB Analysis Resource v2.22 for CD4 epitope predictions. Due to the lack of reliable immunogenicity prediction tools for CD4 peptides on IEDB.org, the IFNEpitope server was used to predict IFN γ -inducing epitopes. IFNEpitope is predominantly designed to predict MHC II peptides with the capacity to induce IFN γ release and has an accuracy of 82.10%. The VaxiJen and AllerTOP servers have been regarded as highly reliable antigenicity and allergenicity prediction servers with an accuracy of 87% and 89%, respectively. For toxicity evaluation, ToxinPred was selected due to its accuracy of 94.50%.

We revealed CD8+ and CD4+ T cell epitopes from SARS-CoV-2 S protein targeting the Omicron (BA.1/B.1.1.529) variant while

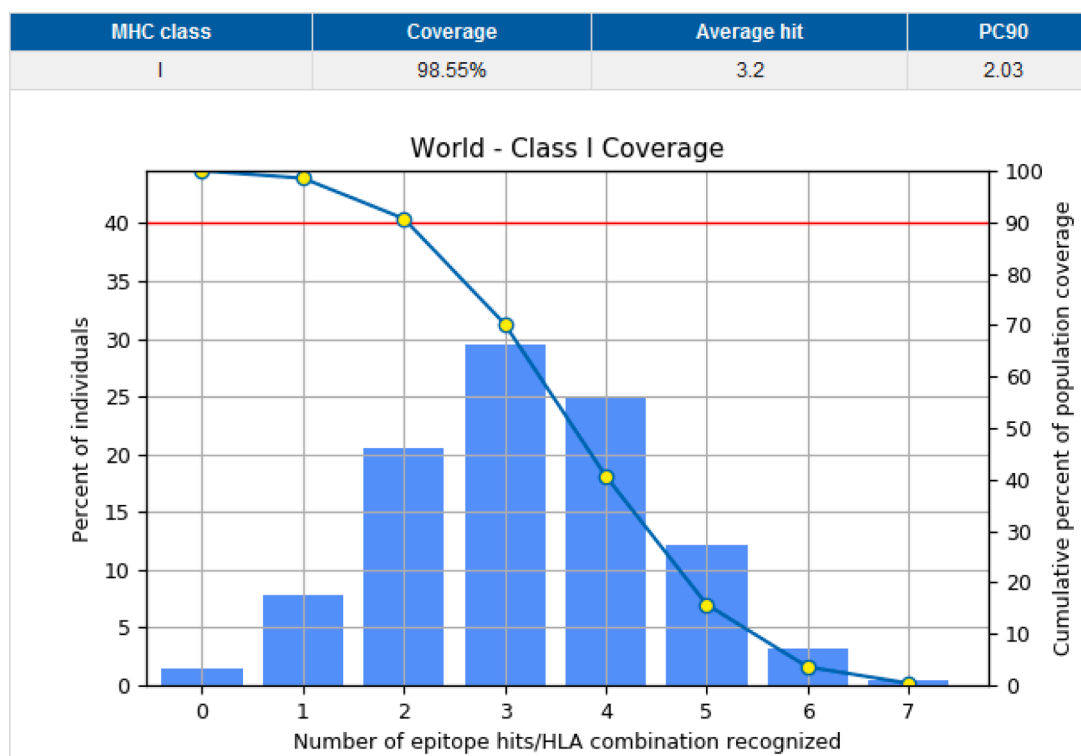


Fig. 3. Population coverage for optimal epitope dataset retrieved from PCOptim. The epitope data accounts for 27 unique HLA alleles which contributes to the high population coverage. Epitope datasets that pass through the clinical checkpoint filters will have fewer unique HLA alleles, and therefore have lower population coverage.

considering many of its significant mutations such as A67V, del69–70, T95I, ins214EPE, S371L, S373P, S375F, Q493K, G496S, Q498R, N501Y, N679K and P681H. We identified eight immunogenic antigenic non-allergenic nontoxic CD8+ T cell epitopes and 11 IFN γ inducing antigenic non-allergenic nontoxic stable CD4+ T cell epitopes that can provide a robust immune response and cover 76.16% and 97.46% of the world population, respectively (Tables 1 and 2). Among the identified epitopes, 5 CD8 peptides (RSYSFRPTY, NLAPFFTFK, YNLAPFFTF, APFFTFKCY and VLYNLAPFF) are located in the RBD region, suggesting their potential to target the impaired interaction between S protein and the neutralizing antibodies. To better understand how our epitopes target Omicron S protein RBD mutations, we obtained the structure of Fab fragment in complex with SARS-CoV-2 S protein (RCSB PDB 6XCM) (Fig. 5). Moreover, our predicted CD8 epitope RSYSFRPTY targets the two important mutations, Q498R and N501Y, correlated with an enhanced binding affinity to the ACE2 and decreased binding affinity to neutralizing antibodies [51]. We predicted RSYSFRPTY to be immunogenic, highly antigenic, non-allergenic and nontoxic with a world population coverage of 49.04% alone, suggesting its importance in epitope-based vaccine design. To further enhance the immune response, we identified two 15-mer peptides (SVLYNLAPFFTFKCY, VLYNLAPFFTFKCYG) that overlap with the identified CD8+ T cell epitopes in the RBD region. Together these epitopes can provide a strong and long-lasting immune response against the Omicron variant.

We used our software tool PCOptim to validate the accuracy of the CD8 top epitope selection. Additionally, the PCOptim tool allowed us to determine the caveats in population coverage and thus identify the populations that may receive less coverage. The PCOptim tool provided us with an optimal set of epitopes that reaches the maximum possible population coverage given our raw dataset. Comparing this optimal dataset to our set of top epitopes revealed some overlap, confirming the success of our selection procedure. A total of three epitopes were found in both the optimal dataset and our top epitope datasets. Out of the 27 most common HLA alleles in the human population, we identified five

alleles that are unaccounted for in the immunogenic, antigenic, non-allergenic, nontoxic and stable epitope dataset (HLA-B*07:02, HLA-B*08:01, HLA-A*31:01, HLA-A*33:01, and HLA-B*51:01). The difference in population coverage between that optimal dataset and our set of 8 CD8 Omicron-related epitopes can be attributed to the five aforementioned HLA alleles. The only way to increase population coverage further is to include new epitopes that are predicted to be strong binders to the desired HLA alleles and pass the clinical checkpoint filters. While PCOptim does not account for the clinical checkpoint parameters we address in this study, it is a useful tool in determining several epitopes that are useful in obtaining high population coverage. The population coverage that we obtained from our top epitope dataset is substantially high, so adding new epitopes will make a very small impact in total population coverage.

Previous studies have shown that SARS-CoV-2 may use mutations on the NTD region of S protein to escape potent polyclonal neutralizing responses [52]. We identified an immunogenic antigenic non-allergenic and nontoxic CD8+ T cell epitope REPEDLPQGF which consists of a novel insertion mutation ins214EPE, not observed in any other SARS-CoV-2 strains [53]. Although, the mutation is believed to be arisen from co-infection of SARS-CoV-2 with HCoV-229E as the two share similar nucleotide sequences, we were unable to detect any common epitopes between the two human-coronavirus strains after running a multiple sequence alignment analysis. To create a robust immune response that targets the NTD region of SARS-CoV-2 Omicron, we identified an additional CD8+ T cell epitope (GVYFASIEK) and two high quality CD4+ T cell epitopes (FLPFFSNVTWFHVIS and LPPFFSNVTWFHVISG) which consist of mutation A67V and a pair of deletions at residues 69 and 70.

We predicted another important immunogenic antigenic non-allergenic nontoxic CD8 peptide (KSHRRARSV) which is located at the furin cleavage site of SARS-CoV-2. The cleavage of the S protein into S1 and S2 is an essential step in viral entry into a host cell [54] and it is believed that the mediation of membrane fusion could be linked to the

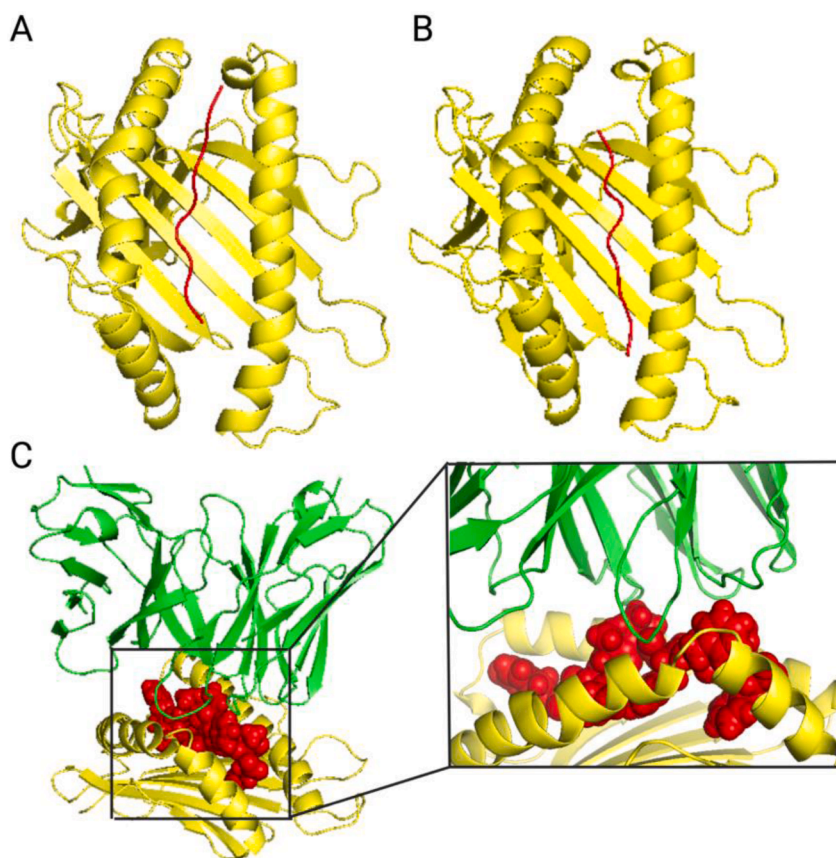


Fig. 4. 3D modeling of two SARS-CoV-2 S peptides in complex with corresponding MHC Class I molecule (A) Cartoon form of CD8+ T cell epitope KSHRRARSV (red) and HLA-A*30:01 allele (yellow). (B) Cartoon form of CD8+ T cell epitope YNLAPFFTF (red) and HLA-A*24:02 allele (yellow). (C) Binding interaction of sphere form of epitope YNLAPFFTF (red) with C1-28 TCR (green) (PDB: 3VXM) and HLA-A*24:02 (yellow).

high virulence of SARS-CoV-2, mainly due to a cluster of mutations at the H655Y, N679K and P681H amino acid sites [18]. KSHRRARSV targets both N679K and P681H and is presented by an antigenic IFN γ inducing non-allergenic nontoxic 15-mer amino acid sequence (KSHRRARSVASQSII) which together can provide a strong protection against B.1.1.529 or any future variants. These epitopes also target the novel circulating sister-lineage BA.2 of Omicron which confers identical mutations observed in B.1.1.529.

The large number of mutations on the RBD of S protein is speculated to help the virus escape neutralizing antibodies from natural and vaccine-induced immunity. Despite the loss of binding affinity to human ACE2 due to mutations such as K417N, Omicron has restored its strong binding to the ACE2 with mutations at residues 493, 496, 498 and 501. Previously we showed how epitopes can target the S protein interaction with neutralizing antibodies (Fig. 5). While completing the manuscript, Omicron specific S protein structure in complex with human ACE2 (PDB 7T9K) became available which we used to visualize our predicted epitopes and their interaction with human ACE2 and Omicron S protein complex. We show that CD8+ T cell epitopes (RSYSFRPTY, NLAPFFTFK, YNLAPFFTF, APFFTFKCY and VLYNLAPFF) on the RBD target the ACE2 binding site (Fig. 6). Together with the predicted CD4+ T cell epitopes, a robust immune response can be created which facilitates the production of antibodies, therefore blocking cell entry. Moreover, NTD and furin cleavage sites are critical to virus attachment to membrane protein and other sites, and we have identified epitopes that can target these sites (Fig. 6). With the emergence of new SARS-CoV-2 variants and their high likelihood to adapt to mutations on the RBD, NTD and furin cleavage sites, we speculate that our identified epitopes in the RBD region can effectively and directly interact with the host ACE2, thereby interfering with the binding affinity. We have already observed

that the BA.2 lineage confers many common mutations in the BA.1/B.1.1.529 variant. We show that CD8 peptide (KSHRRARSV) and CD4 epitope (KSHRRARSVASQSII) are recognized by both Omicron sub-lineages (BA.1 and BA.2) (Table 3). Moreover, these epitopes are located on the S1/S2 cleavage site, associated with increased ACE2 binding affinity and impaired antibody recognition. Therefore, we propose that these epitopes can be used to target the critical site of SARS-CoV-2 S protein in both BA.1 and BA.2 lineages.

Other computational work for prediction of SARS-CoV-2 epitopes have been reported growingly in the literature, including immunoinformatic analysis of sequence homology and prediction of T and B cell epitopes using IEDB tools [55–59]. While multiple studies have previously described immunoreactive CD8 and CD4 peptides in different antigenic regions of SARS-CoV-2 [60,61], exact Omicron-specific T cell epitopes remain yet to be comprehensively determined. In our study, we attempted to identify and validate novel Omicron specific peptides with high population-scale presentation while conferring many clinically relevant properties. A key aspect of our prediction method was to identify epitopes that are common among several SARS-CoV-2 variants and specific to the novel Omicron variant. Furthermore, we presented epitopes with significant mutations in the S protein, which are discussed in several studies to reduce spike-specific T cell reactivity in natural and/or vaccine induced host system with Omicron variant [62,63]. Only three unique Omicron specific CD8 epitopes identified in our work (NSASFSTFK - IEDB epitope ID 1,313,244; GYVFASTK - IEDB epitope ID 1,312,627; YNSASFSTF - IEDB epitope ID 1,397,512) are represented in their unmutated form in the literature. The rest which are not experimentally validated in their unmutated form, are likely creating new T cell epitopes. Another advantage of our study is the prioritization of overlapping CD4+ and CD8+ T cell epitopes that have the potential of

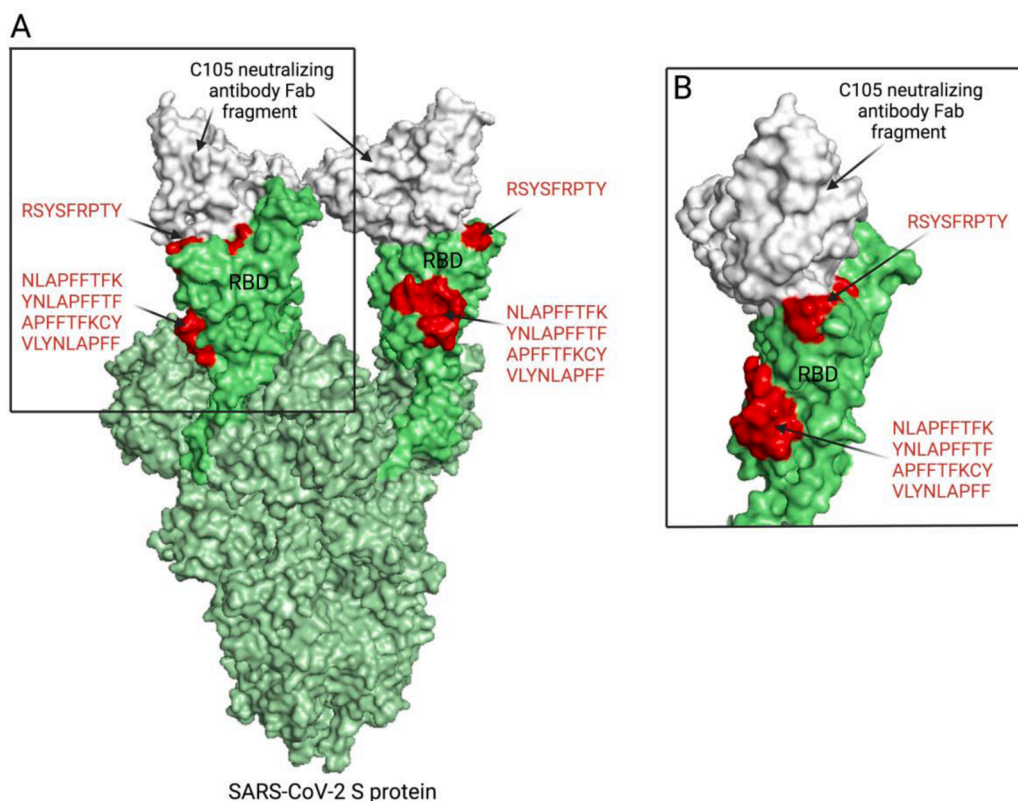


Fig. 5. 3D structure of SARS-CoV-2 S protein in complex with the C105 neutralizing antibody Fab fragment (PDB: 6XCM) with five highlighted CD8+ T cell epitopes specific for Omicron (BA.1/B.1.1.529) variant. (A) SARS-CoV-2 S protein in complex with the C105 neutralizing antibody Fab fragment. Our predicted immunogenic CD8+ T cell epitopes (RSYSFRPTY, NLAPFFTFK, YNLAPFFTF, APFFTFKCY and VLYNLAPFF) are colored in red. (B) Zoomed in version of the complex that focuses on S protein RBD interaction with the neutralizing antibody, highlighting the predicted epitopes.

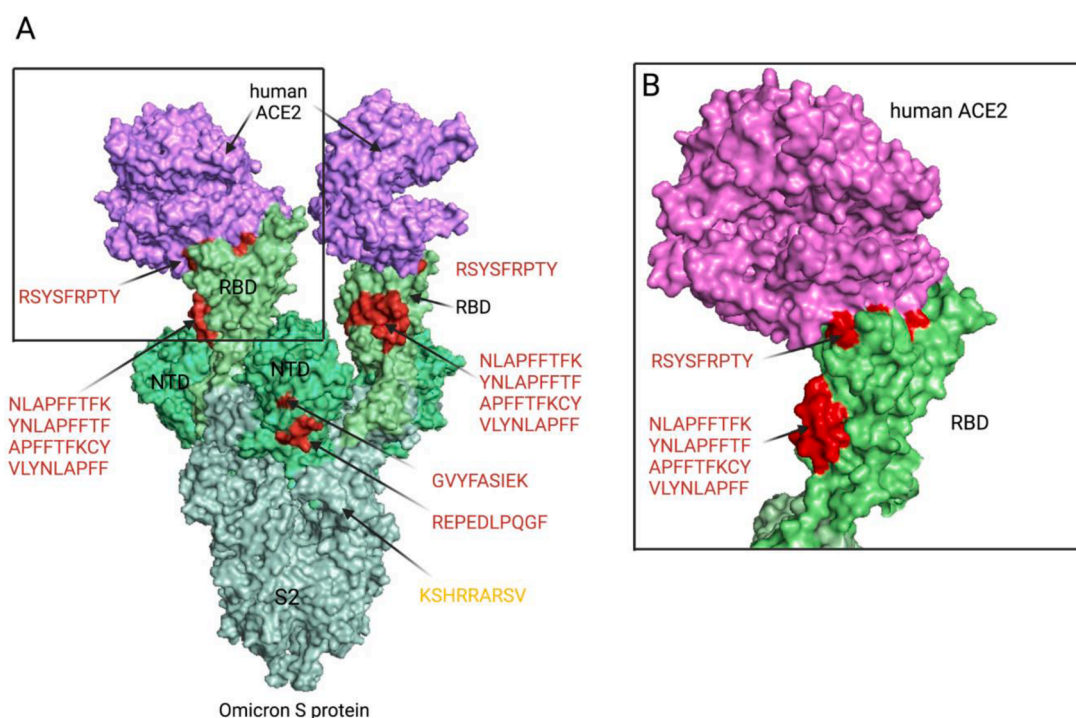


Fig. 6. 3D structure of SARS-CoV-2 Omicron S protein-ACE2 complex (PDB: 7T9K) with eight highlighted CD8+ T cell epitopes. Original Omicron (BA.1/B.1.1.529) specific epitopes are colored in red. Common epitopes among BA.1 and BA.2 variants are colored in yellow. (A) Omicron S protein in complex with human ACE2. Five of our predicted immunogenic CD8+ T cell epitopes (RSYSFRPTY, NLAPFFTFK, YNLAPFFTF, APFFTFKCY and VLYNLAPFF) are within the receptor-binding domain (RBD), targeting the binding site of ACE2. Two immunogenic CD8+ T cell epitopes (GVFYASIEK and REPEDLPQGF) target novel Omicron (BA.1) specific mutations on the NTD. One immunogenic CD8+ T cell epitope (KSHRRARSV) was identified at the furin cleavage site, common for both BA.1 and BA.2 lineages. (B) Zoomed in version of the complex that focuses on Omicron RBD interaction with human ACE2, highlighting the predicted CD8 epitopes on the RBD site.

creating stronger immune responses. Furthermore, our identified epitopes allow for murine vaccine studies using human-relevant peptides in murine MHC haplotypes. Murine models have low-cost maintenance and reflect the clinical signs, viral replication and pathology of SARS-CoV-2 in humans [64], making them useful animal models to assess the safety and efficacy of proposed multi-epitope vaccine candidates in preclinical studies. Predicting our top CD8 epitopes' binding affinity to murine MHC Class I alleles allows for murine vaccine studies using human-relevant peptides. In our analysis, we utilized two mouse strains: C57BL/6 and BALB/CJ, that have been extensively used in SARS-CoV research [65].

The findings through computational analysis indicate that the designed epitopes could be vastly useful in multi-epitope vaccine construct as they have a high probability of being safe and effective in protecting humans from the SARS-CoV-2 virus. To further induce the immunogenicity of the antigens and provide a more robust and long-lasting immune response, we suggest that emulsion adjuvant, like MF59, could be added to the final vaccine construct. Superior to other types of adjuvants, they elicit cell-mediated immune response with improved antigen uptake and ability to promote the migration of activated antigen-presenting cells. Furthermore, MF59 has been used in preclinical trials of coronavirus targeted vaccines [66]. Linkers are also important for vaccines while providing enhanced flexibility, extended protein folding and separation of functional domains. Available studies propose that CD8 and CD4 epitopes could be linked together using AYY linker and GPGPG linker, respectively [67].

Limitations of study

The *in-silico* prediction analysis of the SARS-CoV-2 S protein reported in our study requires further laboratory validation to select the most efficient vaccine candidate.

Conclusion

Recent emergence of the highly mutated Omicron variant of SARS-CoV-2 has disrupted confidence around whether the current vaccines and antibody therapies will provide long-term protection against the novel coronavirus. Subsequently, the possibility of escape from natural and vaccine-induced immunity has prompted an urgent need for new vaccine constructs which target the most concerning mutations on the variant. Using immunoinformatic methods and bioinformatics tools, we identified immunogenic antigenic non-allergenic nontoxic CD8+ T cell epitopes and IFN γ inducing antigenic non-allergenic nontoxic CD4+ T cell epitopes on the RBD, NTD and furin cleavage sites of S protein which target Omicron specific mutations by interfering with S protein interaction to host ACE2. Moreover, our analysis yielded common high-quality CD8+ and CD4+ T cell epitopes across Omicron (BA.1, BA.2) and other circulating SARS-CoV-2 variants including Delta, Alpha, Beta, Gamma, US variants (S protein mutations), and Cluster 5 mink variants with a world population coverage of 84.0% and 96.65%, respectively. We validated the findings using our software PCOptim which enabled us to identify an optimal set of epitopes reaching the maximum possible population coverage. Our *in-silico* epitope prediction together with murine MHC affinity prediction allows scientists to validate the efficacy of the proposed multi-peptide vaccine model through further preclinical studies. The structural validation of the vaccine candidates and the binding affinity to MHC and TCR molecules provide useful insight into immune response mechanism. Because of the rapid development of SARS-CoV-2, we anticipate that the sets of T cell epitopes reported here may be considered as valuable, safe, and efficient approach to control the Omicron variant as well as any future SARS-CoV-2 variants.

CRedit authorship contribution statement

Simone Parn: Methodology, Data curation, Formal analysis,

Investigation, Writing – original draft, Writing – review & editing, Visualization. **Kush Savsani:** Data curation, Formal analysis, Software, Validation, Writing – review & editing. **Sivanesan Dakshanamurthy:** Conceptualization, Methodology, Software, Investigation, Resources, Project administration, Supervision, Writing – original draft, Writing – review & editing, Funding acquisition.

Declaration of Competing Interests

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

Data Availability

Data will be made available on request.

Acknowledgments

We wish to acknowledge the support of the Lombardi Comprehensive Cancer Center, Georgetown University Medical Center, COVID-19 pilot award. This work was supported in part by funding from Georgetown Lombardi's Cancer Research Training and Education Coordination (CRTEC), and the author S.P. and K.S. were part of the GLCCC Undergraduate Summer Research Program.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.immuno.2022.100020](https://doi.org/10.1016/j.immuno.2022.100020).

References

- [1] Classification of omicron (B.1.1.529): SARS-CoV-2 variant of concern. World Health Organization 2021.
- [2] Science Brief: Omicron (B.1.1.529) Variant. Centers for Disease Control and Prevention 2021. Accessed December 19, 2021.
- [3] Cameroni E, Bowen JE, Rosen LE, et al. Broadly neutralizing antibodies overcome SARS-CoV-2 Omicron antigenic shift. Nature 2021. <https://doi.org/10.1038/s41586-021-04386-2>.
- [4] Huang Y, Yang C, Xi Xu, et al. Structural and functional properties of SARS-CoV-2 spike protein: potential antiviral drug development for COVID-19. Acta Pharmacol Sin 2020;41:1141–9. <https://doi.org/10.1038/s41401-020-0485-4>.
- [5] Duan L, et al. The SARS-CoV-2 spike glycoprotein biosynthesis, structure, function, and antigenicity: implications for the design of spike-based vaccine immunogens. Front Immunol 2020. <https://doi.org/10.3389/fimmu.2020.576622>.
- [6] Wang MY, et al. SARS-CoV-2: structure, biology, and structure-based therapeutics development. Front Cell Infect Microbiol 2020. <https://doi.org/10.3389/fcimb.2020.587269>.
- [7] Yu F, Xiang R, Deng X, et al. Receptor-binding domain-specific human neutralizing monoclonal antibodies against SARS-CoV and SARS-CoV-2. Sig Transduct Target Ther. 2020;5:212. <https://doi.org/10.1038/s41392-020-00318-0>.
- [8] Liu L, Iketani S, Guo Y, et al. Striking antibody evasion manifested by the omicron variant of SARS-CoV-2. 2021. Nature 2022. <https://doi.org/10.1038/s41586-021-04388-0>.
- [9] Gao. Ancestral SARS-CoV-2-specific T cells cross-recognize Omicron (B.1.1.529). bioRxiv 2022. <https://doi.org/10.21203/rs.3.rs-1217466/v1>.
- [10] Patronov A, Doytchinova I. T-cell epitope vaccine design by immunoinformatics. Open Biol 2013;12:0139. <https://doi.org/10.1098/rsob.120139>.
- [11] De Groot AS, Moise L, McMurry JA, Martin W. Epitope-Based Immunome-Derived Vaccines: A Strategy for Improved Design and Safety. Clin Appl Immunomics 2008; 2:39–69. https://doi.org/10.1007/978-0-387-79208-8_3.
- [12] Ng OW, et al. Memory T cell responses targeting the SARS coronavirus persist up to 11 years post-infection. Vaccine 2016;34:2008–14.
- [13] Robbiani DF, et al. Convergent antibody responses to SARS-CoV-2 in convalescent individuals. Nature 2020;584:437–42.
- [14] Man S, Newberg MH, Crotzer VL, et al. Definition of a human T cell epitope from influenza A non-structural protein 1 using HLA-A2.1 transgenic mice. Int Immunol 1995;7(4):597–605. <https://doi.org/10.1093/intimm/7.4.597>.
- [15] Shirai M, Arichi T, Nishioka M, Nomura T, Ikeda K, Kawanishi K, et al. CTL responses of HLA-A2.1-transgenic mice specific for hepatitis C viral peptides predict epitopes for CTL of humans carrying HLA-A2.1. J Immunol (Baltimore, Md.: 1950) 1995;154(6):2733–42.

- [16] Vita R, Mahajan S, Overton JA, Dhanda SK, Martini S, Cantrell JR, et al. The Immune Epitope Database (IEDB): 2018 update. *Nucleic Acids Res* 2018. <https://doi.org/10.1093/nar/gky1006>.
- [17] Savsani K, Jabbour G, Dakshanamurthy S. A new epitope selection method: application to design a multi-valent epitope vaccine targeting hras oncogene in squamous cell carcinoma. *Vaccines* 2022;10:63. <https://doi.org/10.3390/vaccines10010063>.
- [18] Hodcroft, E. Variant: 21 K (Omicron). CoVariants. Accessed January 13, 2022.
- [19] BA.2 mutations: SARS-CoV-2 variants of concern and variants under investigation in England. Technical briefing 31. UK Health Security Agency. 2021. Accessed January 25, 2022.
- [20] Nielsen M, Lundegaard C, Blicher T, et al. NetMHCpan, a method for quantitative predictions of peptide binding to any HLA-A and -B locus protein of known sequence. *PLoS One* 2007;2(8):e796. <https://doi.org/10.1371/journal.pone.0000796>.
- [21] Hoof I, Peters B, Sidney J, et al. NetMHCpan, a method for MHC class I binding prediction beyond humans. *Immunogenetics* 2009;61(1):1–13. <https://doi.org/10.1007/s00251-008-0341-z>.
- [22] Nielsen M, Andreatta M, NetMHCpan-3.0; improved prediction of binding to MHC class I molecules integrating information from multiple receptor and peptide length datasets. *Genome Med* 2016;8(1):33. <https://doi.org/10.1186/s13073-016-0288-x>.
- [23] Jurtz V, Paul S, Andreatta M, Marcattili P, Peters B, Nielsen M. NetMHCpan-4.0: Improved peptide-MHC class I interaction predictions integrating eluted ligand and peptide binding affinity data. *J Immunol* 2017;199(9):3360–8. <https://doi.org/10.4049/jimmunol.1700893>.
- [24] Reynisson B, Alvarez B, Paul S, Peters B, Nielsen M. NetMHCpan-4.1 and NetMHCIIpan-4.0: Improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data. *Nucleic Acids Res* 2020;48(W1):W449–54. <https://doi.org/10.1093/nar/gkaa379>.
- [25] Calis JJA, Maybeno M, Greenbaum JA, et al. Properties of MHC class I presented peptides that enhance immunogenicity. *PLoS Comput Biol* 2013;9(10). <https://doi.org/10.1371/journal.pcbi.1003266>.
- [26] Wang P, Sidney J, Kim Y, Sette A, Lund O, Nielsen M, et al. Peptide binding predictions for HLA DR, DP and DQ molecules. *BMC Bioinformatics* 2020;11:568. <https://doi.org/10.1186/1471-2105-11-568>.
- [27] 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(4):e1000048. <https://doi.org/10.1371/journal.pcbi.1000048>.
- [28] Dhanda SK, Vir P, Raghava GPS. Designing of Interferon-Gamma Inducing MHC Class-II Binders. *Biol Direct* 2013;8:30. <https://doi.org/10.1186/1745-6150-8-30>.
- [29] Doytchinova IA, Flower DR. VaxiJen: A server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinformatics* 2007;8:4. <https://doi.org/10.1186/1471-2105-8-4>.
- [30] Dimitrov I, Bangov I, Flower DR, Doytchinova I. AllerTOP v.2- a server for in silico prediction of allergens. *J Mol Model* 2014;20:2278. <https://doi.org/10.1007/s00894-014-2278-5>.
- [31] Gupta S, Kapoor P, Chaudhary K, et al. Silico Approach for Predicting Toxicity of Peptides and Proteins. *PLoS ONE*. 2013;8:e73957. <https://doi.org/10.1371/journal.pone.0073957>.
- [32] Gasteiger E., Hoogland C., Gattiker A., et al. Protein identification and analysis tools on the Expasy server. In: John M. Walker, ed. *The proteomics protocols handbook*, Humana Press; 2005:571–607.
- [33] Bui HH, Sidney J, Dinh K, Southwood S, Newman MJ, Sette A. Predicting population coverage of T-cell epitope-based diagnostics and vaccines. *BMC Bioinformatics* 2006;17:153. <https://doi.org/10.1186/1471-2105-7-153>.
- [34] Rammensee HG, Bachmann J, Emmerich NP, Bachor OA, Stevanović S. SYFPEITHI: database for MHC ligands and peptide motifs. *Immunogenetics* 1999;50(3–4): 213–9. <https://doi.org/10.1007/s002510050595>.
- [35] Wu T, Guan J, Handel A, et al. Quantification of epitope abundance reveals the effect of direct and cross-presentation on influenza CTL responses. *Nat Commun* 2019;10(1):2846. <https://doi.org/10.1038/s41467-019-10661-8>.
- [36] Ikuta Y, Hayashida Y, Hirata S, Irie A, Senju S, Kubo T, et al. Identification of the H2-Kd-restricted cytotoxic T lymphocyte epitopes of a tumor-associated antigen, SPARC, which can stimulate antitumor immunity without causing autoimmune disease in mice. *Cancer Sci* 2009;100(1):132–7. <https://doi.org/10.1111/j.1349-7006.2008.01016.x>.
- [37] Van der Most RG, Sette A, Oseroff C, et al. Analysis of cytotoxic T cell responses to dominant and subdominant epitopes during acute and chronic lymphocytic choriomeningitis virus infection. *J Immunol* 1996;157(12):5543–54.
- [38] Kaur H, Garg A, Raghava GPS. PEPstr: A de novo method for tertiary structure prediction of small bioactive peptides. *Protein Pept Lett* 2007;14(7):626–31. <https://doi.org/10.2174/092986607781483859>.
- [39] London N, Raveh B, Cohen E, Fathi G, Schueler-Furman O. Rosetta FlexPepDock web server—high resolution modeling of peptide-protein interactions. *Nucleic Acids Res* 2011;39:249. <https://doi.org/10.1093/nar/gkr431>.
- [40] van Zundert GCP, Rodrigues JPGLM, Trellet M, Schmitz C, Kastriitis PL, Karaca E, et al. The HADDOCK2.2 webserver: User-friendly integrative modeling of biomolecular complexes. *J Mol Biol* 2016;428(2015):720–5.
- [41] Gowthaman R, Pierce BG. TCRmodel: high resolution modeling of T cell receptors from sequence. *Nucleic Acids Res* 2018;46:W396–401. <https://doi.org/10.1093/nar/gky432>.
- [42] Parn S, Jabbour G, Nguyenkhao V, Dakshanamurthy S. Design of Peptide Vaccine for COVID-19: CD8+ and CD4+ T cell epitopes from SARS-CoV-2 open-reading frame protein variants. *bioRxiv* 2021:461301. <https://doi.org/10.1101/2021.09.21.461301>.
- [43] Jabbour G, Rego S, Nguyenkhao V, Dakshanamurthy S. Design of T cell epitope-based vaccine candidate for SARS-CoV-2 targeting spike and nucleocapsid protein escape variants. *bioRxiv* 2021:459907. <https://doi.org/10.1101/2021.09.11.459907>.
- [44] Zhou D, et al. Evidence of escape of SARS-CoV-2 variant B.1.351 from natural and vaccine-induced sera. *Cell* 2021;189:2348–61. <https://doi.org/10.1016/j.cell.2021.02.037>.
- [45] Wang Z, Schmidt F, Weisblum Y, et al. mRNA vaccine-elicited antibodies to SARS-CoV-2 and circulating variants. *Nature* 2021;592:616–22. <https://doi.org/10.1038/s41586-021-03324-6>.
- [46] Shah M., Woo H.G. Omicron: A heavily mutated SARS-CoV-2 variant exhibits stronger binding to ACE2 and potentially escape approved COVID-19 therapeutic antibodies. *bioRxiv* 2021.12.04.471200. Accessed January 10, 2022. Doi: 10.1101/2021.12.04.471200.
- [47] Andrews N, et al. Effectiveness of COVID-19 vaccines against the Omicron (B.1.1.529) variant of concern. *medRxiv* 2021. <https://doi.org/10.1101/2021.12.14.21267615>.
- [48] Health Individuals DFDiscovery. South Africa's largest private health insurance administrator, releases at-scale, real-world analysis of Omicron outbreak based on 211 000 COVID-19 test results in South Africa, including collaboration with the South Africa. New York: Discovery Ltd. 2021 Accessed on January 11, 2022.
- [49] Thomas SJ, Moreira ED, et al. Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine through 6 Months. *N Engl J Med* 2021;385:1761–73. <https://doi.org/10.1056/NEJMoa2110345>.
- [50] Adhikari UK, Tayebi M, Rahman MM. Immunoinformatics Approach for Epitope-Based Peptide Vaccine Design and Active Site Prediction against Polyprotein of Emerging Oropouche Virus. *J Immunol Res* 2018;2018:6718083. <https://doi.org/10.1155/2018/6718083>.
- [51] Lupala C.S., Ye Y., Chen H., Su X.D., Liu H. Mutations in RBD of SARS-CoV-2 Omicron variant result stronger binding to human ACE2 protein. *bioRxiv*. 2021.12.10.472102. Accessed January 9, 2022. Doi: 10.1101/2021.12.10.472102.
- [52] Andreatta E, et al. SARS-CoV-2 escape from a highly neutralizing COVID-19 convalescent plasma. *PNAS* 2021;118(36). <https://doi.org/10.1073/pnas.2103154118>.
- [53] Venkatakrishnan A.J., Anand P., Lenehan P., Suratekar R., Raghunathan B., Niesen M.J. and Soundararajan V. Omicron variant of SARS-CoV-2 harbors a unique insertion mutation of putative viral or human genomic origin. 2021. Preprint. Accessed on January 5, 2022. Doi: 10.31219/osf.io/f7txy.
- [54] Xia X. Domains and Functions of Spike Protein in Sars-Cov-2 in the Context of Vaccine Design. *Viruses* 2021;13(1):109. <https://doi.org/10.3390/v13010109>.
- [55] Baruah V, Bose S. Immunoinformatics-aided identification of T cell and B cell epitopes in the surface glycoprotein of 2019-nCoV. *J Med Virol*. 2020;92(5): 495–500. <https://doi.org/10.1002/jmv.25698>.
- [56] Grifoni A, Sidney J, Zhang Y, Scheuermann RH, Peters B, Sette A. A Sequence Homology and Bioinformatic Approach Can Predict Candidate Targets for Immune Responses to SARS-CoV-2. *Cell Host Microbe* 2020 Apr 8;27(4):671–80. <https://doi.org/10.1016/j.chom.2020.03.002>.
- [57] Lin L, Ting S, Yufei H, Wendong L, Yubo F, Jing Z. Epitope-based peptide vaccines predicted against novel coronavirus disease caused by SARS-CoV-2. *Virus Res* 2020 Oct 15;288:198082. <https://doi.org/10.1016/j.virusres.2020.198082>.
- [58] Bhatnager R, Bhasin M, Arora J, Dang AS. Epitope based peptide vaccine against SARS-COV2: an immune-informatics approach. *J Biomol Struct Dyn* 2021 Sep;39 (15):5690–705. <https://doi.org/10.1080/07391102.2020.1787227>.
- [59] Rakib A, Sami SA, Mimi NJ, Chowdhury MM, Eva TA, Nainu F, Paul A, Shahriar A, Tareq AM, Emon NU, Chakraborty S, Shil S, Mily SJ, Ben Hadda T, Almalki FA, Emran TB. Immunoinformatics-guided design of an epitope-based vaccine against severe acute respiratory syndrome coronavirus 2 spike glycoprotein. *Comput Biol Med*. 2020 Sep;124:103967. <https://doi.org/10.1016/j.combiomed.2020.103967>.
- [60] Jieping D, Junping P, Minghui Q, et al. Identification of HLA-A2 restricted CD8+ T cell epitopes in SARS-CoV-2 structural proteins. *J Leukocyte Biol* 2021;110(6). <https://doi.org/10.1002/JLB.4MA0621-020R>.
- [61] Tarke A, Sidney J, Kidd CK, Dan JM, Ramirez SI, Yu ED, Mateus J, da Silva Antunes R, Moore E, Rubiro P, Methot N, Phillips E, Mallal S, Frazier A, Rawlings SA, Greenbaum JA, Peters B, Smith DM, Crotty S, Weiskopf D, Grifoni A, Sette A. Comprehensive analysis of T cell immunodominance and immunoprevalence of SARS-CoV-2 epitopes in COVID-19 cases. *Cell Rep Med* 2021 Feb 16;2(2):100204. <https://doi.org/10.1016/j.xcrm.2021.100204>.
- [62] Naranbhai V, Nathan A, Kaseke C, Berrios C, Khatri A, et al. T cell reactivity to the SARS-CoV-2 Omicron variant is preserved in most but not all individuals. *Cell* 2022;185(6):1041–51. <https://doi.org/10.1016/j.cell.2022.01.029>.
- [63] Cao Y, Wang J, Jian F, Xiao T, Song W, Yisimayi A, Huang W, Li Q, Wang P, An R, Wang J, Wang Y, Niu X, Yang S, Liang H, Sun H, Li T, Yu Y, Cui Q, Liu S, Yang X, Du S, Zhang Z, Hao X, Shao F, Jin R, Wang X, Xiao J, Wang Y, Xie XS. Omicron escapes the majority of existing SARS-CoV-2 neutralizing antibodies. *Nature* 2022 Feb;602(7898):657–63. <https://doi.org/10.1038/s41586-021-04385-3>.
- [64] Zhang Y, Huang K, Wang T, et al. SARS-CoV-2 Rapidly Adapts in Aged BALB/c Mice and Induces Typical Pneumonia. *J Virol* 2021;95(11). <https://doi.org/10.1128/JVI.02477-20>.
- [65] Corbett KS, Edwards DK, Leist SR, Abiona OM, Boyoglu-Barnum S, Gillespie RA, Himansu S, Schäfer A, Ziwawo CT, DiPiazza AT, Dinnon KH, Elbashir SM, Shaw CA, Woods A, Fritch EJ, Martinez DR, Bock KW, Minai M, Nagata BM, Hutchinson GB, Wu K, Henry C, Bahl K, Garcia-Dominguez D, Ma L, Renzi I, Kong WP, Schmidt SD, Wang L, Zhang Y, Phung E, Chang LA, Loomis RJ,

Altaras NE, Narayanan E, Metkar M, Presnyak V, Liu C, Louder MK, Shi W, Leung K, Yang ES, West A, Gully KL, Stevens LJ, Wang N, Wrapp D, Doria-Rose NA, Stewart-Jones G, Bennett H, Alvarado GS, Nason MC, Ruckwardt TJ, McLellan JS, Denison MR, Chappell JD, Moore IN, Morabito KM, Mascola JR, Baric RS, Carfi A, Graham BS. SARS-CoV-2 mRNA vaccine design enabled by prototype pathogen

preparedness. *Nature* 2020 Oct;586(7830):567–71. <https://doi.org/10.1038/s41586-020-2622-0>.

[66] Liang Z, Zhu H, Wang X, Jing B, Li Z, Xia X, et al. Adjuvants for coronavirus vaccines. *Front Immunol* 2020;11:589833. <https://doi.org/10.3389/fimmu.2020.589833>.

[67] Dong R, Chu Z, Yu F, Zha Y. Contriving multi-epitope subunit of vaccine for COVID-19: immunoinformatics approaches. *Front Immunol* 2020;11:1784. <https://doi.org/10.3389/fimmu.2020.01784>.