

# Bioinformatic prediction of immunodominant regions in spike protein for early diagnosis of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)

Siqi Zhuang<sup>1</sup>, Lingli Tang<sup>1</sup>, Yufeng Dai<sup>1</sup>, Xiaojing Feng<sup>1</sup>, Yiyuan Fang<sup>1</sup>, Haoneng Tang<sup>1</sup>, Ping Jiang<sup>1</sup>, Xiang Wu<sup>2</sup>, Hezhi Fang<sup>3</sup> and Hongzhi Chen<sup>4</sup>

<sup>1</sup> Department of Laboratory Medicine, The Second Xiangya Hospital, Central South University, Changsha, Hunan, China

<sup>2</sup> Department of Parasitology, Xiangya School of Basic Medicine, Central South University, Changsha, Hunan, China

<sup>3</sup> Key Laboratory of Laboratory Medicine, Ministry of Education, Zhejiang Provincial Key Laboratory of Medical Genetics, College of Laboratory Medicine and Life Sciences, Wenzhou Medical University, Wenzhou, Zhejiang, China

<sup>4</sup> National Clinical Research Center for Metabolic Disease, Key Laboratory of Diabetes Immunology, Ministry of Education, Metabolic Syndrome Research Center, and Department of Metabolism & Endocrinology, The Second Xiangya Hospital, Central South University, Changsha, Hunan, China

## ABSTRACT

**Background:** To contain the pandemics caused by SARS-CoV-2, early detection approaches with high accuracy and accessibility are critical. Generating an antigen-capture based detection system would be an ideal strategy complementing the current methods based on nucleic acids and antibody detection. The spike protein is found on the outside of virus particles and appropriate for antigen detection.

**Methods:** In this study, we utilized bioinformatics approaches to explore the immunodominant fragments on spike protein of SARS-CoV-2.

**Results:** The S1 subunit of spike protein was identified with higher sequence specificity. Three immunodominant fragments, Spike<sub>56-94</sub>, Spike<sub>199-264</sub>, and Spike<sub>577-612</sub>, located at the S1 subunit were finally selected via bioinformatics analysis. The glycosylation sites and high-frequency mutation sites on spike protein were circumvented in the antigen design. All the identified fragments present qualified antigenicity, hydrophilicity, and surface accessibility. A recombinant antigen with a length of 194 amino acids (aa) consisting of the selected immunodominant fragments as well as a universal Th epitope was finally constructed.

**Conclusion:** The recombinant peptide encoded by the construct contains multiple immunodominant epitopes, which is expected to stimulate a strong immune response in mice and generate qualified antibodies for SARS-CoV-2 detection.

Submitted 9 December 2020

Accepted 16 March 2021

Published 8 April 2021

Corresponding author

Hongzhi Chen,  
chenhongzhi2013@csu.edu.cn

Academic editor

Elliot Lefkowitz

Additional Information and  
Declarations can be found on  
page 15

DOI 10.7717/peerj.11232

© Copyright

2021 Zhuang et al.

Distributed under

Creative Commons CC-BY 4.0

OPEN ACCESS

**Subjects** Bioinformatics, Computational Biology, Virology, Immunology, Infectious Diseases  
**Keywords** SARS-CoV-2, Spike protein, Antigen-capture, Immunodominant fragments, COVID-19

## INTRODUCTION

The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is highly contagious and has caused more than one hundred million infection cases and over 2.4 million deaths (<https://www.who.int/>, as of February 15, 2021), posing a huge economic and social burden internationally (*Lan et al., 2020; Shang et al., 2020*). The reports of SARS-CoV-2 reinfection cases suggest that stronger international efforts are required to prevent COVID-19 re-emergence in the future (*Zhan, Deverman & Chan, 2020*). Nevertheless, the possibility of SARS-CoV-2 becoming a seasonal epidemic cannot be excluded (*Shaman & Galanti, 2020*). Even worse, the large number of asymptomatic infections greatly increase the difficulties of epidemic control (*Rothe et al., 2020*). At present, no specific drugs have been developed for SARS-CoV-2, and the effectiveness of the vaccines on the market still needs time to be evaluated. Therefore, early detection and isolation of infected people are still indispensable means to control the spread of the epidemic, which requires accurate, early, economical, and easy-to-operate diagnostic methods (*Yan, Chang & Wang, 2020*).

The real-time reverse transcriptase-polymerase chain reaction (RT-PCR) and antibody-capture serological tests are currently the main diagnostic methods for SARS-CoV-2 (*Ishige et al., 2020*). As the golden standard, RT-PCR is highly reliable (*Bustin & Nolan, 2020; Padoan et al., 2020*). However, the implementation costs and relatively cumbersome operation problems make it a big challenge for large population screening (*Thabet et al., 2020*). The antibody-capture serological test is convenient, but seroconversion generally occurs in the second or third week of illness. Therefore, it is not ideal for the early diagnosis of infection (*Hachim et al., 2020; Liu et al., 2020; Tang et al., 2020*). The antigen-capture test is an alternative diagnostic method that relies on the immunodetection of viral antigens in clinical samples. Accordingly, this method could be applied for the detection of early infection no matter if the patient was asymptomatic or not (*Ohnishi, 2008*). Compared with RT-PCR based detection method, it is relatively inexpensive and can be used at the point-of-care.

Rapid viral antigen detection has been successfully used for diagnosing respiratory viruses such as influenza and respiratory syncytial viruses (*Cazares et al., 2020; Ji et al., 2011; Ohnishi et al., 2005, 2012; Qiu et al., 2005*). The sensitivity and specificity of the antigen-capture detection system depend highly on the antigen employed to generate antibodies (*Ohnishi et al., 2012*). The spike protein is one of the structural proteins of SARS-CoV-2, with the majority located on the outside surface of the viral particles (*Fehr & Perlman, 2015; Kumar et al., 2020; Woo et al., 2005*). It has a 76.4% homology with the spike protein of SARS-CoV. *Sunwoo et al. (2013)* showed that the bi-specific spike protein derived monoclonal antibody system exhibited excellent sensitivity in SARS-CoV detection. The virus infection is initiated by the interaction of spike protein receptor-binding domain (RBD) and angiotensin-converting enzyme 2 (ACE2) on host cells. It is

widely accepted that the spike protein is one of the earliest antigenic proteins recognized by the host immune system (Callebaut, Enjuanes & Pensaert, 1996; Chen et al., 2020c; Gomez et al., 1998; Lu et al., 2004; Sanchez et al., 1999). Nevertheless, the difficulties of using spike protein as an antigen are also obvious. Firstly, it is not easy to express and purify the full-length spike protein (Tan et al., 2004). Besides, the spike protein is highly glycosylated (Kumar et al., 2020) and prone to mutation (Wang et al., 2020a), which may counteract the sensitivity of antigen-capture based detection method. Hence, it is critical to truncating the glycosylation and mutation sites on spike protein as much as possible in antigen design (Meyer, Drosten & Muller, 2014; Tan et al., 2004). A study using the truncated spike protein to detect SARS-CoV achieved a diagnostic sensitivity of >99% and a specificity of 100% (Mu et al., 2008), which suggests that the truncated spike protein of SARS-CoV-2 could also be an appropriate candidate for the early diagnostic testing and screening of SARS-CoV-2. In this study, we analyzed the spike protein via bioinformatics tools to obtain immunodominant fragments. The predicted sequences were joined together as a novel antigen for the immunization of mice and antibody production. Epitopes information presented by this work may aid in developing a promising antigen-capture based detection system in pandemic surveillance and containment.

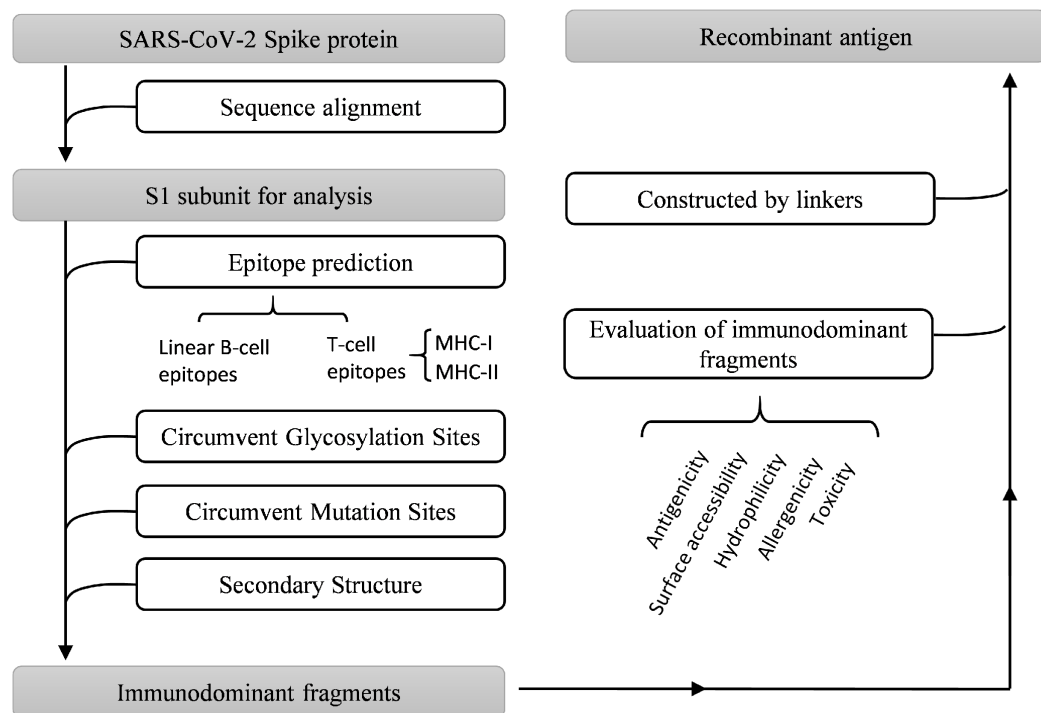
## METHOD

### Data retrieval and sequence alignment

Multiple bioinformatics analysis tools were used in this study, and the flowchart is depicted in Fig. 1. Coronaviruses had four genera composed of alpha-, beta-, gamma- and delta-coronaviruses. Among them, alpha- and beta- genera could infect humans. Seven beta-coronaviruses are known to infect humans (HCoV-229E, HCoV-OC43, HCoV-NL63, HCoV-HKU1, SARS-CoV, MERS-CoV, and SARS-CoV-2) (Kin et al., 2015; Su et al., 2016). We utilized the NCBI database to obtain the sequences of these human-related coronaviruses spike proteins, of which accession numbers were presented in Fig. 2A. The Clustal Omega Server-Multiple Sequence Alignment was used to analyze the sequence similarity. The analysis of the phylogenetic tree was calculated by the same server. In this study, we set parameters of Clustal Omega as default (Sievers et al., 2011). Additionally, we conducted the EMBOSS Needle Server-Pairwise Sequence Alignment (Needleman & Wunsch, 1970) to compare the whole sequence and several major domains between SARS-CoV-2 and SARS-CoV to find out the specific genomic regions on SARS-CoV-2.

### Linear B-cell epitope prediction

Linear B-cell epitopes of the SARS-CoV-2 spike protein were calculated by ABCpred and Bepipred v2.0 servers. For ABCpred, we set a threshold of 0.8 to achieve a specificity of 95.50% and an accuracy of 65.37% for prediction. The window length was set to 16 (the default window length) in this study (Saha & Raghava, 2006). The BepiPred v2.0 combines a hidden Markov model and a propensity scale method. The score threshold for the BepiPred v2.0 was set to 0.5 (the default value) to obtain a specificity of 57.16% and a



**Figure 1** Work flow chart.

Full-size  DOI: 10.7717/peerj.11232/fig-1

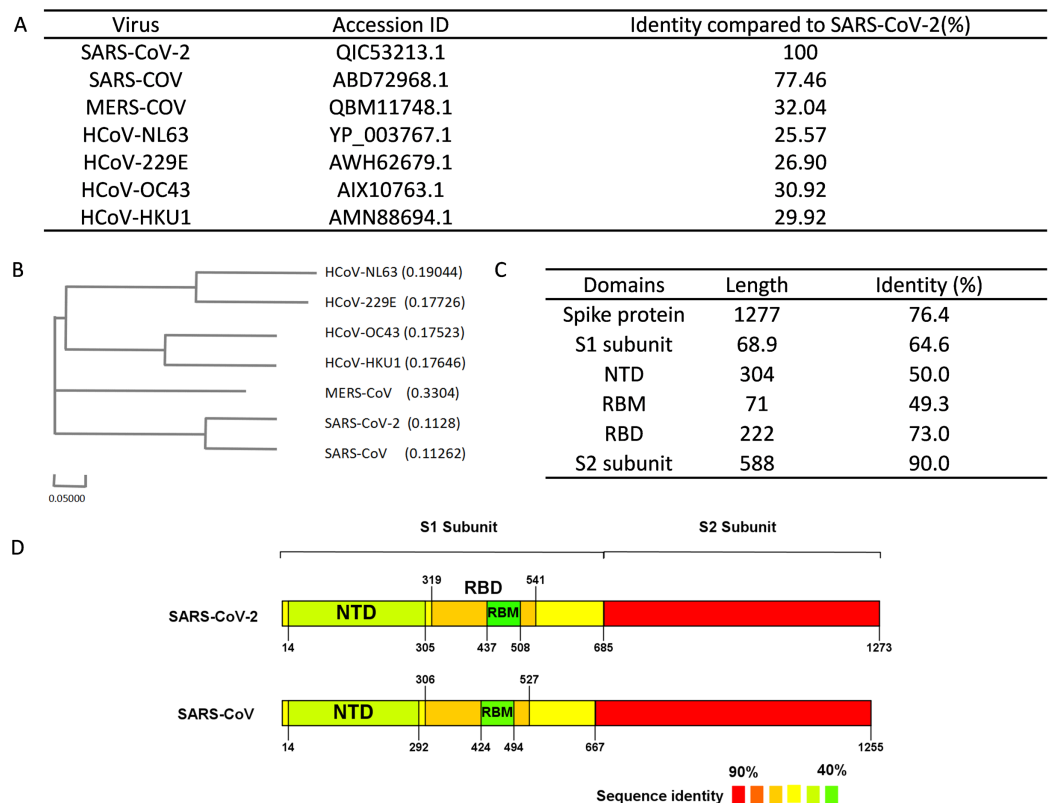
sensitivity of 58.56% (Jespersen *et al.*, 2017). The residues with scores above 0.5 were predicted to be part of an epitope.

### T-cell epitope prediction

The free online service TepiTool server, integrated into the Immune Epitope Database (IEDB), was used to forecast epitopes binding to mice MHC molecules (Paul *et al.*, 2016). Alleles including H-2-Db, H-2-Dd, H-2-Kb, H-2-Kd, H-2-Kk, and H-2-Ld were selected for MHC-I binding epitopes analysis. We checked the “IEDB recommended” option during computation and retained sequences with predicted consensus percentile rank  $\leq 1$  as predicted epitope (Trolle *et al.*, 2015). For MHC-II binding epitopes, alleles including H2-IAb, H2-IAd, and H2-IEd were selected for analysis. As the same as MHC-I binding computation, we chose the “IEDB recommended” option, and peptides with predicted consensus percentile rank  $\leq 10$  were identified as potential epitopes (Wang *et al.*, 2010; Zhang *et al.*, 2012).

### Profiling and evaluation of selected fragments

The secondary structure of the SARS-CoV-2 spike protein (PDB ID: 6VSB chain B) was calculated by the PyMOL molecular graphics system using the SSP algorithm. PyMOL (<http://www.pymol.org>) is a python-based tool, which is widely used for visualization of macromolecules, such as SARS-CoV-2 spike protein in the current study (Yuan *et al.*, 2016). Vaxijen2.0 server was utilized to analyze the antigenicity of epitopes and selected fragments. A default threshold of 0.4 was set and the prediction accuracy is between 70% and 89% (Doytchinova & Flower, 2007). The hydrophilicity of the selected fragment was



**Figure 2** Sequence alignment results of spike protein. (A) Accession IDs and sequence identities of selected coronavirus spike protein. (B) Phylogenetic tree of spike proteins among selected coronavirus. (C) Sequence identity of major domains in spike protein between SARS-CoV-2 and SARS-CoV. (D) Sequence identity of domains in SARS-CoV-2 and SARS-CoV reflected by colors. From red to green, the color changing represents the sequence identity from high to low.

Full-size [DOI: 10.7717/peerj.11232/fig-2](https://doi.org/10.7717/peerj.11232/fig-2)

analyzed by the online server ProtScale (Wilkins *et al.*, 1999). Surface accessibility of predicted fragments was evaluated by NetsurfP, an online server calculating the surface accessibility and secondary structure of amino acid sequence (Petersen *et al.*, 2009). Critical features such as allergenicity and toxicity were evaluated by online server AllerTOP v2.0 (Dimitrov *et al.*, 2014) and ToxinPred (Gupta *et al.*, 2013). In addition, we utilized IEDB ([www.iedb.org](http://www.iedb.org)) to search the selected fragments and epitopes to clarify whether these peptides have been experimentally verified (Vita *et al.*, 2019). Protein sequence BLAST was performed to evaluate the possibility of cross-reactivity with other mouse protein sequences (Altschul *et al.*, 1997).

## RESULTS

### Sequence alignment of spike protein in different coronaviruses

We performed sequence alignment to determine the evolutionary relationships between SARS-CoV-2 and other beta-coronaviruses that could infect humans. According to the results of sequence alignment (Figs. 2A and 2B), SARS-CoV is the closest virus to SARS-CoV-2 among the seven HCoVs, exhibiting a 77.46% sequence identity. To better

understand the divergence of spike protein sequences between SARS-CoV-2 and SARS-CoV, we further analyzed the sequences of main domains. Results showed that the S2 subunit was the most conserved domain with a 90.0% identity. RBM and NTD domains, which were located in the S1 subunit, exhibited 49.3% and 50.0% identity respectively (Figs. 2C and 2D). Hence, we chose the S1 subunit (amino acid 1-685) for the subsequent bioinformatics analysis given their high specificity.

### Linear B-cell epitope prediction of S1 subunit in SARS-CoV-2 spike protein

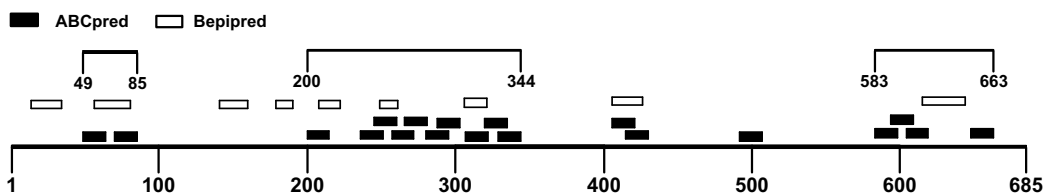
The B-cell epitope is a surface accessible cluster of amino acids, which could be recognized by secreted antibodies or B-cell receptors and elicit humoral immune response (Getzoff *et al.*, 1988). The immunodominant fragments should contain high-quality linear B-cell epitopes to stimulate antibody production effectively. The sequence of the SARS-CoV-2 S1 subunit was evaluated via ABCpred and BepiPred v2.0. A total of 31 peptides were identified by the ABCpred algorithm (Table S1). For the BepiPred v2.0 server, 14 epitopes were forecasted (Table S2). After antigenicity evaluation, 19 and 9 potential linear B-cell epitopes predicted by the ABCpred server and BepiPred v2.0 server were obtained respectively (Table 1). The peptides predicted by both bioinformatics programs are more likely to be an epitope recognized *in vivo*. After mapping the positions of peptides identified by these servers, 3 regions containing predicted epitopes were obtained. These regions could be preliminarily considered as candidates for immunodominant fragments (Fig. 3; Table 2).

### Murine T-cell epitope prediction of S1 subunit in SARS-CoV-2 spike protein

Though B cells are responsible for producing antibodies, humoral immunity is heavily dependent on the activation of T cells (Cho *et al.*, 2019a). Helper T cells (Th) recognize antigen peptides presented by MHC-II molecules and facilitate the humoral immune response (Cho *et al.*, 2019b; Mahon *et al.*, 1995). During humoral immune responses, antigen-activated T cells could provide help in many aspects including directing antibody class switching and guiding the differentiation of antibody-secreting plasma cells as well as the properties of the B-cell antigen receptor (Cho *et al.*, 2019a; Paus *et al.*, 2006; Shulman *et al.*, 2014). Therefore, the immunodominant fragments containing T-cell epitopes could offer essential help to powerful antibody production. The S1 subunit was selected for the prediction of T-cell epitopes. We utilized the TepiTool server to forecast MHC-I and MHC-II binding epitopes. A total of 35 MHC-I binding epitopes was predicted (Table S3), and 27 peptides were identified as MHC-II binding epitopes (Table S4). The antigenicity of these peptides was calculated via Vaxijen 2.0 server (Table 3). Combined with the MHC-II epitopes prediction results, the candidate immunodominant fragments were adjusted (Fig. 4). Compared with the preliminary candidate immunodominant fragments screened according to the linear B-cell epitope prediction, we added the Spike<sub>14-34</sub> fragment into consideration because it contains a linear

**Table 1** Linear B-cell epitopes predicted by ABCpred and BepiPred v2.0 with antigenicity score exceed the threshold value.

Tools	Position	Sequence	Length	Antigenicity (cut off $\geq 0.4$ )
ABCpred	583-598	EILDITPCSFGGVSVI	16	1.3971
	406-421	EVRQIAPGQTGKIADY	16	1.3837
	415-430	TGKIADYNYKLPDDFT	16	0.9642
	648-663	GCLIGAHEVNNSEYCD	16	0.848
	288-303	AVDCALDPLSEKCTL	16	0.7905
	604-619	TSNQVAVLYQDVNCTE	16	0.7593
	307-322	TVEKGIYQTSNFRVQP	16	0.6733
	200-215	YFKIYSKHTPINLVRD	16	0.657
	257-272	GWTAGAAAYVGYLQP	16	0.621
	329-344	FPNITNLCPFGEVFNA	16	0.6058
	245-260	HRSYLTPGDSSSGWTA	16	0.6017
	280-295	NENGTITDAVDCALDP	16	0.5804
	49-64	HSTQDLFLPFFSNVTW	16	0.5305
	492-507	LQSYGFQPTNGVGYQP	16	0.5258
	70-85	VSGTNGTKRFDNPVLP	16	0.5162
	236-251	TRFQTLALHRSYLTP	16	0.5115
	266-281	YVGYLQPRTFLLKYNE	16	0.5108
	594-609	GVSVITPGTNTSNQVA	16	0.4651
	320-335	VQPTEIVRFPNITNL	16	0.4454
	Bepipred v2.0	179-190	LEGKQGNFKNLR	12
404-426		GDEVQRQIAPGQTGKIADYNYKLP	23	1.1017
14-34		QCVNLTTRTQLPPAYTNSFTR	21	0.7594
56-81		LPFFSNVTWFHAIHVSNGTKRFDN	26	0.6041
208-222		TPINLVRDLPQGFS	15	0.5531
141-160		LGVYYHKNKSWMESEFRVY	20	0.5308
249-261		LTPGDSSSGWTAG	13	0.495
306-321		FTVEKGIYQTSNFRVQ	16	0.4361
615-644		VNCTEVPVAIHADQLTPTWRVYSTGNSNVFQ	30	0.4259

**Figure 3** Preliminary immunodominant fragments based on B-cell epitope prediction results. The black squares represent epitopes predicted by ABCpred server, the black frames represent epitopes predicted by Bepipred v2.0 server, and the black lines with numbers on both ends represent the preliminary candidate immunodominant fragments. [Full-size !\[\]\(86257f54800c9844bc7e863bea396fba\_img.jpg\) DOI: 10.7717/peerj.11232/fig-3](https://doi.org/10.7717/peerj.11232/fig-3)

**Table 2** Details of epitopes in the preliminary immunodominant fragments selected according to linear B-cell epitope prediction results.

Regions	Epitope predicted by ABCpred			Epitope predicted by Bepipred v2.0		
	Position	Sequence	Antigenicity	Position	Sequence	Antigenicity
49-85	49-64	HSTQDLFLPFFSNVTW	0.5305	56-81	LPFFSNVTWFHAIHVSNGTKRFDN	0.6041
	70-85	VSGTNGTKRFDNPVLP	0.5162			
200-344	200-215	YFKIYSKHTPINLVRD	0.6570	208-222	TPINLVRDLPQGFS	0.5531
	236-251	TRFQTLALHRSYLTP	0.5115	249-261	LTPGDSSSGWTAG	0.4950
	245-260	HRSYLTPGDSSSGWTA	0.6017			
	257-272	GWTAGAAAYVGYLQP	0.6210			
	266-281	YVGYLQPRTFLLKYNE	0.5108			
	280-295	NENGTITDAVDCALDP	0.5804			
	288-303	AVDCALDPLSETKCTL	0.7905	306-321	FTVEKGIYQTSNFRVQ	0.4361
	307-322	TVEKGIYQTSNFRVQP	0.6733			
	320-335	VQPTESIVRFPNITNL	0.4454			
	329-344	FPNITNLCPFGEVFNA	0.6058			
583-663	415-430	TGKIADYNYKLPDDFT	0.9642			
	583-598	EILDITPCSFGGVSVI	1.3971	615-644	VNCTEVPVAIHADQLTPTWRVYSTGSNVFQ	0.4259
	594-609	GVSVITPGTNTSNQVA	0.4651			
	604-619	TSNQVAVLYQDVNCTE	0.7593			
	648-663	GCLIGAEHVNNSEYCD	0.8480			

B epitope and an MHC-II binding epitope, both of which had high antigenicity scores (Table 4).

### Immunodominant fragments refinement according to the glycosylation site distribution, mutation site distribution, and secondary structure

A profile of 24 glycosylation sites of SARS-CoV-2 spike protein has been reported (Shajahan *et al.*, 2020). Since glycans could hinder the recognition of antigens by shielding the residues (Walls *et al.*, 2019), protein glycosylation would affect the performance of antigen detection. Thus, glycosylation sites should be circumvented when selecting the immunodominant fragments. According to the study of Shajahan *et al.* (2020), 15 glycosylation sites were located in the S1 subunit of the spike protein. Hence, the fragments in this study were adjusted to Spike<sub>14-34</sub>, Spike<sub>49-101</sub>, Spike<sub>199-261</sub>, and Spike<sub>583-620</sub>. To retain antigenicity of the epitopes, the final identified fragments only contained 3 glycosylation sites which should have a minimum effect on antigen recognition.

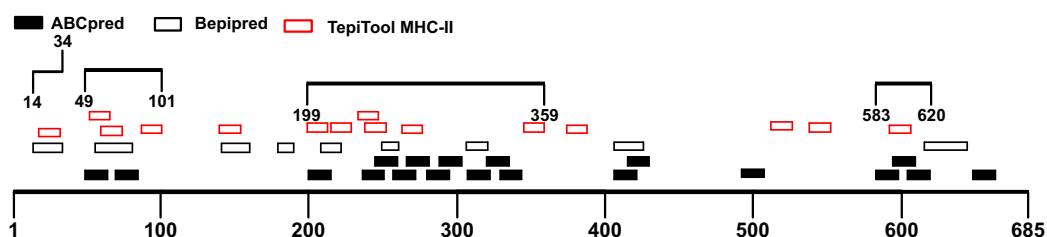
Rapid transmission of COVID-19 provides the SARS-CoV-2 with substantial opportunities for natural selection and mutations. To ensure the stability of the detection method, the immunodominant fragments were modified to avoid high-frequency mutation sites (Wang *et al.*, 2020b). Spike<sub>14-34</sub> were excluded for containing four high-frequency mutation sites. Fragment Spike<sub>49-101</sub> was adjusted to Spike<sub>56-92</sub>, and fragment Spike<sub>583-620</sub> was adjusted to Spike<sub>583-609</sub>. By adjusting the fragments, we avoided



**Table 3** MHC-II and MHC-I binding epitopes predicted by TepiTool server with antigenicity score exceed threshold value.

Type	Position	Sequence	Length	Allele	Core (smm-align)	Core (nn-align)	Percentile Rank	Antigenicity (cut off $\geq 0.4$ )
MHC-II binding	538-552	CVNFNFNGLTGTGVL	15	H2-IAb	FNFNGLTGT	FNFNGLTGT	8.55	1.3281
	374-388	FSTFKCYGVSPKLN	15	H2-IAb	FKCYGVSP	YGVSPKLN	6.45	1.0042
	199-213	GYFKIYSKHTPINLV	15	H2-Iab	KIYSKHTPI	YSKHTPINL	6.9	0.9278
	18-32	LTTRTQLPPAYTNSF	15	H2-IAb	TRTQLPPAY	TRTQLPPAY	9.9	0.79
	60-74	SNVTWFHAIHVSGTN	15	H2-IAb	VTWFHAIHV	TWFHAIHVS	9.1	0.7044
	263-277	AAYYVGYLQPRTFLL	15	H2-IAb	VGYLQPRTF	VGYLQPRTF	8.75	0.6073
	592-606	FGGVSVITPGTNTSN	15	H2-IAb	VITPGTNTS	VSVITPGTN	6	0.5825
	238-252	FQTLALHRSYLTGP	15	H2-IEd	TLLALHRSY	TLLALHRSY	9.85	0.5789
	345-359	TRFASVYAWNRKRIS	15	H2-IAb	FASVYAWNR	YAWNRKRIS	7.45	0.4963
	215-229	DLPQGFSALEPLVDL	15	H2-IAb	FSALEPLVD	FSALEPLVD	6.05	0.4812
	140-154	FLGVYYHKNNKSWME	15	H2-IEd	GVYYHKNNK	YYHKNNKSW	6.4	0.4793
	512-526	VLSFELLHAPATVCG	15	H2-IAb	FELLHAPAT	FELLHAPAT	2.9	0.4784
	87-101	NDGVYFASTEKSNI	15	H2-Iab	YFASTEKSN	VYFASTEKS	6.85	0.4277
	52-66	QDLFLPFFSNVTWFH	15	H2-IAb	FLPFFSNVT	FLPFFSNVT	2.95	0.4159
233-247	INITRFQTLLALHRS	15	H2-IAAd	ITRFQTLA	ITRFQTLA	1.9	0.4118	
MHC-I binding	643-651	FQTRAGCLI	9	H-2-Kk			0.6	1.7332
	612-620	YQDVNCTEV	9	H-2-Db			0.4	1.6172
	539-547	VNFNFNGLT	9	H-2-Kb			0.47	1.5069
	503-511	VGYQPYRVV	9	H-2-Kb			0.47	1.4383
	379-387	CYGVSPTKL	9	H-2-Kd			0.3	1.4263
	16-24	VNLTRTQL	9	H-2-Kb			0.86	1.3468
	510-518	VVLSFELL	9	H-2-Kb			0.43	1.0909
	202-210	KIYSKHTPI	9	H-2-Kb			0.27	0.7455
	168-176	FEYVSQPFL	9	H-2-Kk			0.5	0.6324
	268-276	GYLQPRTF	9	H-2-Kd			0.2	0.6082
	505-513	YQPYRVVVL	9	H-2-Dd			0.3	0.5964
	488-496	CYFPLQSYG	9	H-2-Kd			0.64	0.578
	215-223	DLPQGFSA	9	H-2-Dd			0.69	0.5622
	342-350	FNATRFASV	9	H-2-Kb			0.56	0.5609
	84-92	LPFNDGVYF	9	H-2-Ld			0.21	0.5593
	484-492	EGFNCYFPL	9	H-2-Kb			0.84	0.5453
	62-70	VTWFHAIHV	9	H-2-Kb			0.61	0.5426
	489-497	YFPLQSYGF	9	H-2-Dd			0.8	0.5107
	350-358	VYAWNRKRI	9	H-2-Kd			0.7	0.5003
	60-68	SNVTWFHAI	9	H-2-Kb			0.82	0.4892
262-270	AAAYYVGYL	9	H-2-Kb			0.98	0.4605	

in a total of 8 high-frequency mutation sites (L5F, L18F, T29I, R21K/T, H49Y, L54F, S98F, D614G). The mainly mutant sites on the recent emergent highly infectious variants (including B.1.1.7, B.1.351, and P.1), such as N501Y, D614G, E484K, Y144del, K417N, and A570D were also not included in our fragments. The adjusted fragments contain none of



**Figure 4** Adjusted candidate immunodominant fragments according to MHC-II T-cell epitope prediction results. The black squares represent epitopes predicted by ABCpred server, and the black frames represent epitopes predicted by Bepipred v2.0 server. The red frames denote MHC-II binding epitopes. The black lines with numbers on both ends represent the adjusted candidate fragments.

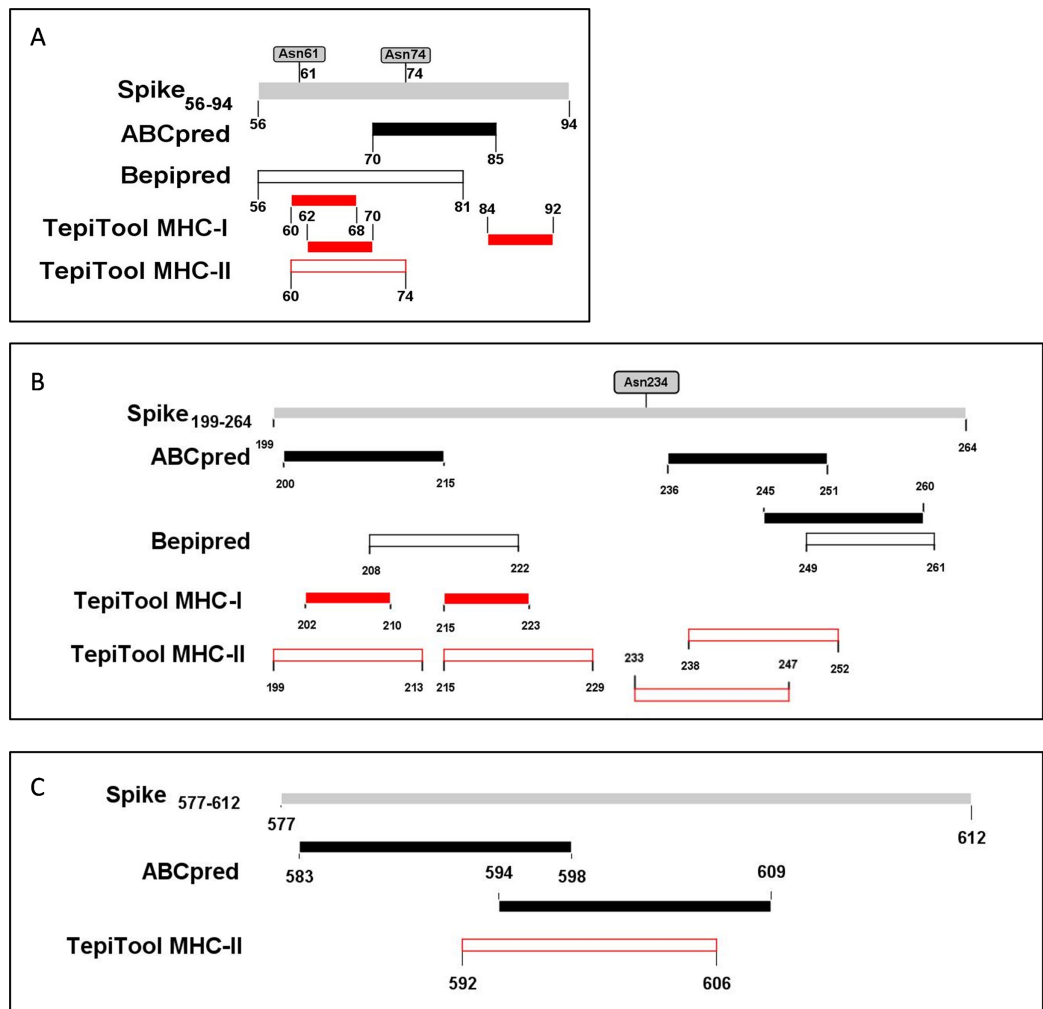
Full-size [DOI: 10.7717/peerj.11232/fig-4](https://doi.org/10.7717/peerj.11232/fig-4)

**Table 4** Details of candidate immunodominant fragments adjusted according to the MHC-II binding T-cell epitopes prediction results.

Regions	Linear B-cell epitopes				MHC-II binding epitopes		
	Tools	Position	Sequence	Antigenicity	Position	Sequence	Antigenicity
14-34	Bepipred v2.0	14-34	QCVNLTTRTQLPPAYTNSFTR	0.7594	18-32	LTTRTQLPPAYTNSF	0.7900
49-101	Bepipred v2.0	56-81	LPFFSNVTWFHAIHVSGTNGTKRFDN	0.6041	52-66	QDLFLPFFSNVTWFH	0.4159
	ABCpred	49-64	HSTQDLFLPFFSNVTW	0.5305	60-74	SNVTWFHAIHVSGTN	0.7044
199-359	ABCpred	70-85	VSGTNGTKRFDNPVLP	0.5162	87-101	NDGVYFASTEKSNII	0.4277
	Bepipred v2.0	208-222	TPINLVRDLPPQGFS	0.5531	199-213	GYFKIYKHTPINLV	0.9278
583-620	ABCpred	249-261	LTPGDSSSGWTAG	0.4950			
		306-321	FTVEKGIYQTSNFRVQ	0.4361			
		200-215	YFKIYKHTPINLVRD	0.6570	215-229	DLPQGFSALEPLVDL	0.4812
		236-251	TRFQTLALHRSYLTP	0.5115	233-247	INITRFQTLALHRS	0.4118
		245-260	HRSYLTPGDSSSGWTA	0.6017	238-252	FQTLALHRSYLTPG	0.5789
		257-272	GWTAGAAAYVGYLQP	0.6210	263-277	AAYVGYLQPRTFLL	0.6073
		266-281	YVGYLQPRTFLLKYNE	0.5108			
		280-295	NENGTITDAVDCALDP	0.5804			
		288-303	AVDCALDPLSETKCTL	0.7905			
		307-322	TVEKGIYQTSNFRVQP	0.6733			
		320-335	VQPTESIVRFPNITNL	0.4454			
583-620	ABCpred	329-344	FPNITNLCPFGEVFN	0.6058	345-359	TRFASVYAWNRKRIS	0.4963
		583-598	EILDITPCSGGVSVI	1.3971	592-606	FGGVSIVTPTNTSN	0.5825
		594-609	GVSIVTPTNTSNQVA	0.4651			
604-619	TSNQVAVLYQDVNCTE	0.7593					

the above high-frequency mutation sites, which might avoid the impact of mutations on detection performance and improve the detection efficiency in the future (*Li et al., 2020; Tegally et al., 2020*).

The PyMOL was used to present the secondary structure of the spike protein (PDB ID: 6VSB) (*Fig. S1*). To keep the integrity of the secondary structure of the selected fragments, we extended the N- and C-ends with 2~5 residues, and the immunodominant fragments were finally adjusted to Spike<sub>56-94</sub>, Spike<sub>199-264</sub>, and Spike<sub>577-612</sub>. The epitopes



**Figure 5** The epitopes and glycosylation sites on the selected immunodominant fragments. (A–C) Present the predicted epitopes and glycosylation sites on fragment Spike56-94, Spike199-264 and Spike577-612 respectively. The black squares represent epitopes predicted by ABCpred server, and the black frames represent epitopes predicted by Bepipred v2.0 server. The red squares represent MHC-I binding epitopes, and the red frames represent MHC-II binding epitopes. The gray squares mean occupied glycosylation sites contained in the selected fragments.

Full-size DOI: [10.7717/peerj.11232/fig-5](https://doi.org/10.7717/peerj.11232/fig-5)

and potential glycosylation sites contained in the selected immunodominant fragments were displayed in [Fig. 5](#).

### Profiling, evaluation, and visualization of selected immunodominant fragments

To further evaluate the antibody binding potentiality of these antigenic regions, the key features of the selected fragments such as antigenicity, hydrophilicity, surface accessibility, toxicity, and allergenicity were analyzed and presented ([Table 5](#)). The hydrophilicity and surface accessibility of the spike protein subunit 1 were calculated. The selected fragments of interest were submitted for computation of antigenicity, toxicity, and allergenicity. Three fragments presented relatively moderate hydrophilicity and surface

**Table 5 Significant features of the selected immunodominant fragments.** The sequences marked as bold and italic in the table represent amino acids with hydrophilicity and surface accessibility respectively.

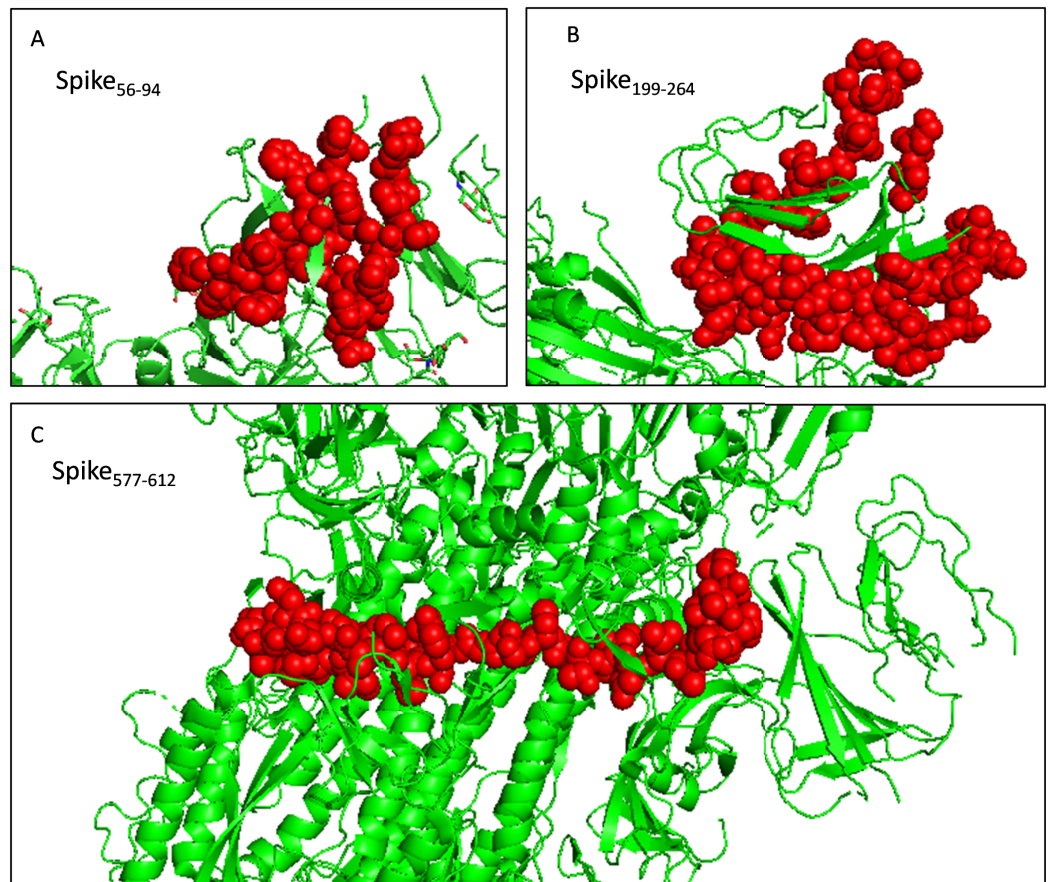
Fragments	Spike <sub>56-94</sub>	Spike <sub>199-264</sub>	Spike <sub>577-612</sub>
Length(aa)	39	66	36
Sequence	LPFFSNVTWFHAIHVSGTNGTKR FDNPVLPFNDGVYFAS	GYFKIYSKHTPINLVRDLPQGFSAL EPLVDLPIGINITRFQTLALHRSYL TPGDSSSGWTAGAAA	RDPQTLEILDITPCSFGGV SVITPGTNTSNQVAVLY
Antigenicity	0.4590	0.5774	0.9127
Domain	S1(NTD)	S1(NTD)	S1
Hydrophilicity fragments	LPFFSNVTWFHAIHVSGTNGTKR <b>FDNPVLPFNDGVYFAS</b>	<b>GYFKIYSKHTPINLVRDLPQGF</b> SALEPLVDLPIGINITRFQTL ALHRSYLTPGDSSSGWTAGAAA	<b>RDPQTLEILDITPCSFGGVSVIT</b> PGTNTSNQVAVLY
Surface Accessibility fragments	LPFFSNVTWFHAIHVSGTNGTKRF <b>DNPVLPFNDGVYFAS</b>	GYFKIYSK <b>HTPINLVRDLPQGFSAL</b> <b>EPLVDLPIGINITRFQTLALHRS</b> <b>YLTPGDSSSGWTAGAAA</b>	<b>RDPQTLEILDITPCSFGGVSVIT</b> <b>PGTNTSNQVAVLY</b>
Toxicity	Non-toxin	Non-toxin	Non-toxin
Allergenicity	non-allergen	non-allergen	probable allergen

accessibility. The proportion of hydrophilic amino acids in the selected fragments Spike<sub>56-94</sub>, Spike<sub>199-264</sub>, Spike<sub>577-612</sub> are 48.72%, 45.45%, 33.33% respectively. The surface accessibility of these fragments calculated by the online server was shown in Table 5.

The toxicity of the selected fragments was examined and no fragment was predicted to be toxic. The allergenicity was assessed and only fragment Spike<sub>577-612</sub> was predicted to be a probable allergen. Attention should be paid to monitor potential allergic reactions when injecting the recombinant protein into mice. And the selected fragments were presented as the sphere in the trimer structure (Fig. 6). Next, we scanned the selected fragments utilizing the IEDB database to determine whether they were experimentally tested. The results showed that Spike<sub>200-215</sub> (IEDB ID: 1330367) and Spike<sub>238-252</sub> (IEDB ID: 1329417) were identified experimentally as HLA class II epitope in SARS-CoV-2. Spike<sub>84-92</sub> (IEDB ID: 1321049) and Spike<sub>202-210</sub> (IEDB ID: 1319559) have been experimentally proved as HLA-B epitopes (Table S5). These findings enhanced the credibility of the current in silico analysis. The fragments identified would have a strong capacity in stimulating powerful antibody production.

### Immunodominant fragments based recombinant antigen design

Three immunodominant fragments embody several linear B-cell epitopes, MHC-I binding, and MHC-II binding T-cell epitopes were selected. As a universal Th epitope, the PAN DR epitope [PADRE(AKFVAAWTLKAAA)] was added into the construction aiming to boost helper T cell activity (Alexander et al., 2000; Ghaffari-Nazari et al., 2015). (GGGGS)<sub>n</sub> is a widely used flexible linker with the function of segmenting protein fragments, maintaining protein conformation, preserving biological activity, and promoting protein expression (Chen, Zaro & Shen, 2013). Finally, we combined the fragments and a PADRE epitope by linker peptide (GGGGS)<sub>2</sub> and (GGGGS)<sub>3</sub> (Chen, Zaro & Shen, 2013) (Fig. 7). The predicted antigenicity of the final construct (194 aa)



**Figure 6** Selected immunodominant fragments presented as spheres in the trimer structure of spike protein viewed by PyMOL. Selected fragments were presented as red spheres, green cartoons denote unselected sections. (A, B, and C) denote fragments Spike<sub>56-94</sub>, Spike<sub>199-264</sub>, and Spike<sub>577-612</sub> respectively.

Full-size DOI: 10.7717/peerj.11232/fig-6

PADRE	(GSSSS) <sub>2</sub>	Spike <sub>56-94</sub>	(GSSSS) <sub>3</sub>	Spike <sub>199-264</sub>	(GSSSS) <sub>3</sub>	Spike <sub>577-612</sub>
AKFVAAW		LPFFSNVTWFHAIHVS		GYFKIYSKHTPINLVR		RDPQTL EILDITPCS
TLKAAA		NGTKRFDNPVLPFND		VDLPIGINITRFQTL		FGGVS VITPGTNTSN
		FAS		SSGWTAGAAA		QVAVLY

**Figure 7** A schematic diagram of recombinant peptide composed of selected fragments and a PADRE epitope.

Full-size DOI: 10.7717/peerj.11232/fig-7

was 0.5690 (Table 6). A protein BLAST for the final construct was conducted to evaluate the possibility of cross-reactivity. The BLAST result suggested that, except for the SARS-CoV-2 spike protein, no protein would cross-react with the construct (Raw data in the Supplemental Files), which indicated that our fragments possess good specificity.

## DISCUSSION

In this study, the immunodominant fragments within the S1 subunit of the SARS-CoV-2 spike protein were explored. The final construct consists of three immunodominant fragments Spike<sub>56-94</sub>, Spike<sub>199-264</sub>, Spike<sub>577-612</sub>, and a PADRE epitope. The recombinant

**Table 6** The structure and antigenicity of final recombinant peptides.

Final construct	PAN DR + (GGGS) <sub>2</sub> + Spike <sub>56-94</sub> +(GGGS) <sub>3</sub> + Spike <sub>199-264</sub> +(GGGS) <sub>3</sub> + Spike <sub>577-612</sub>
Sequence	AKFVAAWTLKAAAGGGSGGGSLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASGGGGSGGGSGGGSGYFKIYSKHTPINLVRDLPQGF SALEPLVDLPIGINITRFQTLALHRSYLTGPDSSSGWTAGAAAGGGSGGGSGGGSRDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLY
Antigenicity	0.5690

antigen will be used to immunize mice to generate qualified antibody which could be applied for developing an antigen-capture based detection system.

The antibody-based antigen capturing method is user-friendly, time-saving, and economical. Thus, it is an ideal complementary detection strategy especially for early diagnosis and large population screening. The monoclonal antibodies against SARS-CoV have been successfully applied in the immunological antigen-detection of SARS-CoV (*Ohnishi, 2008*). Accordingly, we explored the immunodominant fragments on the spike protein of SARS-CoV-2, which would provide aid in developing an accurate and fast antigen-capture based early detection system for SARS-CoV-2.

We selected the S1 subunit for immunodominant fragments screening after divergence analysis. It had been reported that an S1 antigen-based assay of SARS-CoV could capture the virus as soon as the infection occurs (*Sunwoo et al., 2013*). Jong-Hwan Lee et al. designed a method that could seize and detect spike protein S1 subunit of SARS-CoV-2 using ACE2 receptor and S1-mAb (*Lee et al., 2021*). This finding suggests that it is appropriate to use the S1 subunit for specific and early diagnosis of SARS-CoV-2. Three immunodominant fragments (Spike<sub>56-94</sub>, Spike<sub>199-264</sub>, and Spike<sub>577-612</sub>) were identified in the present study. These sequences will be joined to construct recombinant peptides in the next step. Instead of using inactivated full-length spike protein, we designed a novel recombinant protein construct that increased sequence specificity as well as circumvented mutation sites and glycosylation sites. As the antigen design is based on bioinformatics study, the exact ability of the selected fragments to produce qualified antibodies for virus detection has yet to be determined by experiments.

Noticeably, the spike protein of SARS-CoV-2 is heavily glycosylated. Glycans could shield epitopes during antibody recognition, which may interfere with the detection of viral proteins (*Shajahan et al., 2020*). About 17 N-glycosylation sites along with two O-glycosylation sites were found occupied in the spike protein of SARS-CoV-2 (*Shajahan et al., 2020*). We circumnavigated most glycosylation sites when selecting immunodominant fragments. The three selected fragments in this study only contain 3 glycosylation sites. In case these glycosylation sites impede the diagnostic performance, an additional deglycosylation step with N-glycanase should be applied for the test specimens (*Dermani et al., 2019*), which is a simple and efficient method for deglycosylation (*Hirani, Bernasconi & Rasmussen, 1987; Huang et al., 2015; Lattová et al., 2016; Zheng, Bantog & Bayer, 2011*). Alternatively, an eukaryotic expressing system could be employed to mimic the antigen presented in human cells.

Though coronaviruses can find and repair errors during the replication process (*Wang et al., 2020b*), the SARS-CoV-2 genome still presents a large number of mutations.

Mutations could not only help virus slip past our immune defense, but also spoil the efficiency of diagnostic tests (*Chen et al., 2020b*). In this study, we circumvented high-frequency mutation sites when selecting antigen fragments. In addition, our fragments also avoided RBD regions which are prone to mutation (*Chen et al., 2020b*). The construct finally built contained no high-frequency mutation.

To date, several studies using predictive algorithms to analyze SARS-CoV-2 have been reported (*Alam et al., 2020; Behmard et al., 2020; Can et al., 2020; Chen et al., 2020a; Dong et al., 2020; Poran et al., 2020; Saha, Ghosh & Burra, 2021; Sohail et al., 2021*). However, most of these bioinformatics analyses against SARS-CoV-2 intended to develop effective vaccines to prevent infection and the identified sequences possess high homology with other viruses, especially SARS-CoV (*Bhattacharya et al., 2020; Chen et al., 2020a; Robson, 2020*). On the contrary, the fragments suitable for diagnosis should be unique when compared with other species to ensure the specificity of detection. Therefore, the results obtained from vaccine studies are not ideal for virus detection. In this study, attention was paid to the sequences with high variability, hence the immunodominant fragments identified are more specific. Distinct from vaccine studies, murine MHC alleles were selected in epitopes prediction in this study, so that the designed antigen could trigger a strong humoral immune response in mice. Furthermore, glycosylated sites and recently identified high-frequency mutation sites were deliberately avoided during the screening process to eliminate their potential adverse impact.

In silico analysis has been widely used to mine and identify various pathogens as well as epitopes prediction (*Kiyotani et al., 2020; Liò & Goldman, 2004; Qin et al., 2003; Robson, 2020; Shen et al., 2003*). In this study, identified fragments were further scanned in the IEDB database, and found four peptides contained in the sequences were experimentally validated epitopes (*Table S5*), which reinforced the conclusion of the present study. In the following studies, we will immunize Balb/c mice with the designed antigen to generate mAbs which could be utilized for SARS-CoV-2 diagnosis after evaluating their sensitivity, specificity, and other related properties.

## CONCLUSION

Through bioinformatics analysis, three immunodominant fragments were identified in the present study. After connected by flexible linkers, we acquired a final recombinant peptide with 194 residues. It was predicted to possess high antigenicity and specificity for SARS-CoV-2. Our next move is to express and purify the recombinant protein in a suitable expression system, followed by immunizing the mice with purified immunogen to obtain specific antibodies. The present study would provide aid in developing an antigen-capture based detection system.

## ADDITIONAL INFORMATION AND DECLARATIONS

### Funding

This work was supported by the emergency and special scientific program of COVID-19 epidemic of Changsha, China (grant number: kq2001038) and the National Science

Foundation of China (grant number: 81970746). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

### **Grant Disclosures**

The following grant information was disclosed by the authors:

COVID-19 epidemic of Changsha, China: kq2001038.

National Science Foundation of China: 81970746.

### **Competing Interests**

The authors declare that they have no competing interests.

### **Author Contributions**

- Siqi Zhuang conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.
- Lingli Tang performed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
- Yufeng Dai performed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
- Xiaojing Feng performed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
- Yiyuan Fang performed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
- Haoneng Tang performed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
- Ping Jiang performed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
- Xiang Wu performed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
- Hezhi Fang performed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
- Hongzhi Chen conceived and designed the experiments, performed the experiments, authored or reviewed drafts of the paper, and approved the final draft.

### **Data Availability**

The following information was supplied regarding data availability:

The raw measurements are available in the [Supplemental File](#).

### **Supplemental Information**

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj.11232#supplemental-information>.



## REFERENCES

- Alam A, Khan A, Imam N, Siddiqui M, Waseem M, Malik M, Ishrat R. 2020. Design of an epitope-based peptide vaccine against the SARS-CoV-2: a vaccine-informatics approach. *Briefings in Bioinformatics* **22**(2):1309–1323 DOI [10.1093/bib/bbaa340](https://doi.org/10.1093/bib/bbaa340).
- Alexander J, Del Guercio MF, Maewal A, Qiao L, Fikes J, Chesnut RW, Paulson J, Bundle DR, DeFrees S, Sette A. 2000. Linear PADRE T helper epitope and carbohydrate B cell epitope conjugates induce specific high titer IgG antibody responses. *Journal of Immunology* **164**(3):1625–1633 DOI [10.4049/jimmunol.164.3.1625](https://doi.org/10.4049/jimmunol.164.3.1625).
- Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* **25**(17):3389–3402 DOI [10.1093/nar/25.17.3389](https://doi.org/10.1093/nar/25.17.3389).
- Behmard E, Soleymani B, Najafi A, Barzegari E. 2020. Immunoinformatic design of a COVID-19 subunit vaccine using entire structural immunogenic epitopes of SARS-CoV-2. *Scientific Reports* **10**(1):20864 DOI [10.1038/s41598-020-77547-4](https://doi.org/10.1038/s41598-020-77547-4).
- Bhattacharya M, Sharma AR, Patra P, Ghosh P, Sharma G, Patra BC, Lee SS, Chakraborty C. 2020. Development of epitope-based peptide vaccine against novel coronavirus 2019 (SARS-COV-2): immunoinformatics approach. *Journal of Medical Virology* **92**(6):618–631 DOI [10.1002/jmv.25736](https://doi.org/10.1002/jmv.25736).
- Bustin SA, Nolan T. 2020. RT-qPCR testing of SARS-CoV-2: a primer. *International Journal of Molecular Sciences* **21**(8):3004 DOI [10.3390/ijms21083004](https://doi.org/10.3390/ijms21083004).
- Callebaut P, Enjuanes L, Pensaert M. 1996. An adenovirus recombinant expressing the spike glycoprotein of porcine respiratory coronavirus is immunogenic in swine. *Journal of General Virology* **7**(Pt 2):309–313 DOI [10.1099/0022-1317-77-2-309](https://doi.org/10.1099/0022-1317-77-2-309).
- Can H, Köseoğlu AE, Erkunt Alak S, Güvendi M, Döşkaya M, Karakavuk M, Gürüz AY, Ün C. 2020. In silico discovery of antigenic proteins and epitopes of SARS-CoV-2 for the development of a vaccine or a diagnostic approach for COVID-19. *Scientific Reports* **10**(1):22387 DOI [10.1038/s41598-020-79645-9](https://doi.org/10.1038/s41598-020-79645-9).
- Cazares LH, Chaerkady R, Samuel Weng SH, Boo CC, Cimbro R, Hsu HE, Rajan S, Dall'Acqua W, Clarke L, Ren K, McTamney P, Kallewaard-LeLay N, Ghaedi M, Ikeda Y, Hess S. 2020. Development of a parallel reaction monitoring mass spectrometry assay for the detection of SARS-CoV-2 spike glycoprotein and nucleoprotein. *Analytical Chemistry* **92**(20):13813–13821 DOI [10.1021/acs.analchem.0c02288](https://doi.org/10.1021/acs.analchem.0c02288).
- Chen HZ, Tang LL, Yu XL, Zhou J, Chang YF, Wu X. 2020a. Bioinformatics analysis of epitope-based vaccine design against the novel SARS-CoV-2. *Infectious Diseases of Poverty* **9**(1):88 DOI [10.1186/s40249-020-00713-3](https://doi.org/10.1186/s40249-020-00713-3).
- Chen J, Wang R, Wang M, Wei GW. 2020b. Mutations strengthened SARS-CoV-2 infectivity. *Journal of Molecular Biology* **432**(19):5212–5226 DOI [10.1016/j.jmb.2020.07.009](https://doi.org/10.1016/j.jmb.2020.07.009).
- Chen J, Zhu H, Horby PW, Wang Q, Zhou J, Jiang H, Liu L, Zhang T, Zhang Y, Chen X, Deng X, Nikolay B, Wang W, Cauchemez S, Guan Y, Uyeki TM, Yu H. 2020c. Specificity, kinetics and longevity of antibody responses to avian influenza A(H7N9) virus infection in humans. *Journal of Infection* **80**(3):310–319 DOI [10.1016/j.jinf.2019.11.024](https://doi.org/10.1016/j.jinf.2019.11.024).
- Chen X, Zaro JL, Shen WC. 2013. Fusion protein linkers: property, design and functionality. *Advanced Drug Delivery Reviews* **65**(10):1357–1369 DOI [10.1016/j.addr.2012.09.039](https://doi.org/10.1016/j.addr.2012.09.039).
- Cho S, Raybuck A, Blagih J, Kemboi E, Haase V, Jones R, Boothby M. 2019a. Hypoxia-inducible factors in CD4 T cells promote metabolism, switch cytokine secretion, and T cell help in humoral immunity. *Proceedings of the National Academy of Sciences* **116**(18):8975–8984 DOI [10.1073/pnas.1811702116](https://doi.org/10.1073/pnas.1811702116).

- Cho SH, Raybuck AL, Blagih J, Kemboi E, Haase VH, Jones RG, Boothby MR. 2019b.** Hypoxia-inducible factors in CD4(+) T cells promote metabolism, switch cytokine secretion, and T cell help in humoral immunity. *Proceedings of the National Academy of Sciences* **116(18)**:8975–8984 DOI [10.1073/pnas.1811702116](https://doi.org/10.1073/pnas.1811702116).
- Dermani FK, Samadi P, Rahmani G, Kohlan AK, Najafi R. 2019.** PD-1/PD-L1 immune checkpoint: Potential target for cancer therapy. *Journal of Cellular Physiology* **234(2)**:1313–1325 DOI [10.1002/jcp.27172](https://doi.org/10.1002/jcp.27172).
- Dimitrov I, Bangov I, Flower DR, Doytchinova I. 2014.** AllerTOP v.2—a server for in silico prediction of allergens. *Journal of Molecular Modeling* **20(6)**:2278 DOI [10.1007/s00894-014-2278-5](https://doi.org/10.1007/s00894-014-2278-5).
- Dong R, Chu Z, Yu F, Zha Y. 2020.** Contriving multi-epitope subunit of vaccine for COVID-19: immunoinformatics approaches. *Frontiers in immunology* **11**:1784 DOI [10.3389/fimmu.2020.01784](https://doi.org/10.3389/fimmu.2020.01784).
- Doytchinova IA, Flower DR. 2007.** Vaxijen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinformatics* **8(1)**:4 DOI [10.1186/1471-2105-8-4](https://doi.org/10.1186/1471-2105-8-4).
- Fehr AR, Perlman S. 2015.** Coronaviruses: an overview of their replication and pathogenesis. *Methods in Molecular Biology* **1282(Pt 4)**:1–23 DOI [10.1007/978-1-4939-2438-7\\_1](https://doi.org/10.1007/978-1-4939-2438-7_1).
- Getzoff ED, Tainer JA, Lerner RA, Geysen HM. 1988.** The chemistry and mechanism of antibody binding to protein antigens. *Tumor Immunology* **43(9)**:1–98 DOI [10.1016/s0065-2776\(08\)60363-6](https://doi.org/10.1016/s0065-2776(08)60363-6).
- Ghaffari-Nazari H, Tavakkol-Afshari J, Jaafari MR, Tahaghoghi-Hajghorbani S, Masoumi E, Jalali SA. 2015.** Improving multi-epitope long peptide vaccine potency by using a strategy that enhances CD4+ T help in BALB/c Mice. *PLOS ONE* **10(11)**:e0142563 DOI [10.1371/journal.pone.0142563](https://doi.org/10.1371/journal.pone.0142563).
- Gomez N, Carrillo C, Salinas J, Parra F, Borca MV, Escribano JM. 1998.** Expression of immunogenic glycoprotein S polypeptides from transmissible gastroenteritis coronavirus in transgenic plants. *Virology* **249(2)**:352–358 DOI [10.1006/viro.1998.9315](https://doi.org/10.1006/viro.1998.9315).
- Gupta S, Kapoor P, Chaudhary K, Gautam A, Kumar R, Open Source Drug Discovery C, Raghava GP. 2013.** In silico approach for predicting toxicity of peptides and proteins. *PLOS ONE* **8(9)**:e73957 DOI [10.1371/journal.pone.0073957](https://doi.org/10.1371/journal.pone.0073957).
- Hachim A, Kavian N, Cohen CA, Chin AWH, Chu DKW, Mok CKP, Tsang OTY, Yeung YC, Perera R, Poon LLM, Peiris JSM, Valkenburg SA. 2020.** ORF8 and ORF3b antibodies are accurate serological markers of early and late SARS-CoV-2 infection. *Nature Immunology* **21(10)**:1293–1301 DOI [10.1038/s41590-020-0773-7](https://doi.org/10.1038/s41590-020-0773-7).
- Hirani S, Bernasconi RJ, Rasmussen JR. 1987.** Use of N-glycanase to release asparagine-linked oligosaccharides for structural analysis. *Analytical Biochemistry* **162(2)**:485–492 DOI [10.1016/0003-2697\(87\)90424-6](https://doi.org/10.1016/0003-2697(87)90424-6).
- Huang J, Wan H, Yao Y, Li J, Cheng K, Mao J, Chen J, Wang Y, Qin H, Zhang W, Ye M, Zou H. 2015.** Highly efficient release of glycopeptides from hydrazide beads by hydroxylamine assisted PNGase F deglycosylation for N-Glycoproteome analysis. *Analytical Chemistry* **87(20)**:10199–10204 DOI [10.1021/acs.analchem.5b02669](https://doi.org/10.1021/acs.analchem.5b02669).
- Ishige T, Murata S, Taniguchi T, Miyabe A, Kitamura K, Kawasaki K, Nishimura M, Igari H, Matsushita K. 2020.** Highly sensitive detection of SARS-CoV-2 RNA by multiplex rRT-PCR for molecular diagnosis of COVID-19 by clinical laboratories. *Clinica Chimica Acta* **507**:139–142 DOI [10.1016/j.cca.2020.04.023](https://doi.org/10.1016/j.cca.2020.04.023).

- Jespersen MC, Peters B, Nielsen M, Marcatili P. 2017.** BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. *Nucleic Acids Research* **45**(W1):W24–W29 DOI [10.1093/nar/gkx346](https://doi.org/10.1093/nar/gkx346).
- Ji Y, Guo W, Zhao L, Li H, Lu G, Wang Z, Wang G, Liu C, Xiang W. 2011.** Development of an antigen-capture ELISA for the detection of equine influenza virus nucleoprotein. *Journal of Virological Methods* **175**(1):120–124 DOI [10.1016/j.jviromet.2011.04.016](https://doi.org/10.1016/j.jviromet.2011.04.016).
- Kin N, Miszczak F, Gouilh MA, Vabret A, Consortium E. 2015.** Genomic analysis of 15 human coronaviruses OC43 (HCoV-OC43s) circulating in France from 2001 to 2013 reveals a high intra-specific diversity with new recombinant genotypes. *Viruses* **7**(5):2358–2377 DOI [10.3390/v7052358](https://doi.org/10.3390/v7052358).
- Kiyotani K, Toyoshima Y, Nemoto K, Nakamura Y. 2020.** Bioinformatic prediction of potential T cell epitopes for SARS-Cov-2. *Journal of Human Genetics* **65**(7):569–575 DOI [10.1038/s10038-020-0771-5](https://doi.org/10.1038/s10038-020-0771-5).
- Kumar S, Maurya VK, Prasad AK, Bhatt MLB, Saxena SK. 2020.** Structural, glycosylation and antigenic variation between 2019 novel coronavirus (2019-nCoV) and SARS coronavirus (SARS-CoV). *Virusdisease* **31**(1):13–21 DOI [10.1007/s13337-020-00571-5](https://doi.org/10.1007/s13337-020-00571-5).
- Lan J, Ge J, Yu J, Shan S, Zhou H, Fan S, Zhang Q, Shi X, Wang Q, Zhang L, Wang X. 2020.** Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* **581**(7807):215–220 DOI [10.1038/s41586-020-2180-5](https://doi.org/10.1038/s41586-020-2180-5).
- Lattová E, Bryant J, Skříčková J, Zdráhal Z, Popovič M. 2016.** Efficient procedure for N-Glycan analyses and detection of Endo H-Like activity in human tumor specimens. *Journal of Proteome Research* **15**(8):2777–2786 DOI [10.1021/acs.jproteome.6b00346](https://doi.org/10.1021/acs.jproteome.6b00346).
- Lee JH, Choi M, Jung Y, Lee SK, Lee CS, Kim J, Kim J, Kim NH, Kim BT, Kim HG. 2021.** A novel rapid detection for SARS-CoV-2 spike 1 antigens using human angiotensin converting enzyme 2 (ACE2). *Biosensors & Bioelectronics* **171**(7477):112715 DOI [10.1016/j.bios.2020.112715](https://doi.org/10.1016/j.bios.2020.112715).
- Li Q, Wu J, Nie J, Zhang L, Hao H, Liu S, Zhao C, Zhang Q, Liu H, Nie L, Qin H, Wang M, Lu Q, Li X, Sun Q, Liu J, Zhang L, Li X, Huang W, Wang Y. 2020.** The impact of mutations in SARS-CoV-2 spike on viral infectivity and antigenicity. *Cell* **182**(5):1284–1294.e1289 DOI [10.1016/j.cell.2020.07.012](https://doi.org/10.1016/j.cell.2020.07.012).
- Liò P, Goldman N. 2004.** Phylogenomics and bioinformatics of SARS-CoV. *Trends in Microbiology* **12**(3):106–111 DOI [10.1016/j.tim.2004.01.005](https://doi.org/10.1016/j.tim.2004.01.005).
- Liu W, Liu L, Kou G, Zheng Y, Ding Y, Ni W, Wang Q, Tan L, Wu W, Tang S, Xiong Z, Zheng S. 2020.** Evaluation of Nucleocapsid and Spike Protein-based ELISAs for detecting antibodies against SARS-CoV-2. *J Clin Microbiol.* doi [10.1128/JCM.00461-20](https://doi.org/10.1128/JCM.00461-20).
- Lu L, Manopo I, Leung BP, Chng HH, Ling AE, Chee LL, Ooi EE, Chan SW, Kwang J. 2004.** Immunological characterization of the spike protein of the severe acute respiratory syndrome coronavirus. *Journal of Clinical Microbiology* **42**(4):1570–1576 DOI [10.1128/JCM.42.4.1570-1576.2004](https://doi.org/10.1128/JCM.42.4.1570-1576.2004).
- Mahon BP, Katrak K, Nomoto A, Macadam AJ, Minor PD, Mills KH. 1995.** Poliovirus-specific CD4+ Th1 clones with both cytotoxic and helper activity mediate protective humoral immunity against a lethal poliovirus infection in transgenic mice expressing the human poliovirus receptor. *Journal of Experimental Medicine* **181**(4):1285–1292 DOI [10.1084/jem.181.4.1285](https://doi.org/10.1084/jem.181.4.1285).
- Meyer B, Drosten C, Muller MA. 2014.** Serological assays for emerging coronaviruses: challenges and pitfalls. *Virus Research* **194**(2):175–183 DOI [10.1016/j.virusres.2014.03.018](https://doi.org/10.1016/j.virusres.2014.03.018).

- Mu F, Niu D, Mu J, He B, Han W, Fan B, Huang S, Qiu Y, You B, Chen W. 2008. The expression and antigenicity of a truncated spike-nucleocapsid fusion protein of severe acute respiratory syndrome-associated coronavirus. *BMC Microbiology* 8(1):207 DOI 10.1186/1471-2180-8-207.
- Needleman SB, Wunsch CD. 1970. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *Journal of Molecular Biology* 48(3):443–453 DOI 10.1016/0022-2836(70)90057-4.
- Ohnishi K. 2008. Establishment and characterization of monoclonal antibodies against SARS coronavirus. *Methods in Molecular Biology* 454:191–203 DOI 10.1007/978-1-59745-181-9\_15.
- Ohnishi K, Sakaguchi M, Kaji T, Akagawa K, Taniyama T, Kasai M, Tsunetsugu-Yokota Y, Oshima M, Yamamoto K, Takasuka N, Hashimoto S, Ato M, Fujii H, Takahashi Y, Morikawa S, Ishii K, Sata T, Takagi H, Itamura S, Odagiri T, Miyamura T, Kurane I, Tashiro M, Kurata T, Yoshikura H, Takemori T. 2005. Immunological detection of severe acute respiratory syndrome coronavirus by monoclonal antibodies. *Japanese Journal of Infectious Diseases* 58:88–94.
- Ohnishi K, Takahashi Y, Kono N, Nakajima N, Mizukoshi F, Misawa S, Yamamoto T, Mitsuki YY, Fu S, Hirayama N, Ohshima M, Ato M, Kageyama T, Odagiri T, Tashiro M, Kobayashi K, Itamura S, Tsunetsugu-Yokota Y. 2012. Newly established monoclonal antibodies for immunological detection of H5N1 influenza virus. *Japanese Journal of Infectious Diseases* 65(5):19–27 DOI 10.7883/yoken.65.442.
- Padoan A, Cosma C, Sciacovelli L, Faggian D, Plebani M. 2020. Analytical performances of a chemiluminescence immunoassay for SARS-CoV-2 IgM/IgG and antibody kinetics. *Clinical Chemistry and Laboratory Medicine* 58(7):1081–1088 DOI 10.1515/cclm-2020-0443.
- Paul S, Sidney J, Sette A, Peters B. 2016. TepiTool: a pipeline for computational prediction of T cell epitope candidates. *Current Protocols in Immunology* 114(1):163 DOI 10.1002/cpim.12.
- Paus D, Phan T, Chan T, Gardam S, Basten A, Brink R. 2006. Antigen recognition strength regulates the choice between extrafollicular plasma cell and germinal center B cell differentiation. *Journal of Experimental Medicine* 203(4):1081–1091 DOI 10.1084/jem.20060087.
- Petersen B, Petersen TN, Andersen P, Nielsen M, Lundegaard C. 2009. A generic method for assignment of reliability scores applied to solvent accessibility predictions. *BMC Structural Biology* 9(1):51 DOI 10.1186/1472-6807-9-51.
- Poran A, Harjanto D, Malloy M, Arieta C, Rothenberg D, Lenkala D, Van Buuren M, Addona T, Rooney M, Srinivasan L, Gaynor R. 2020. Sequence-based prediction of SARS-CoV-2 vaccine targets using a mass spectrometry-based bioinformatics predictor identifies immunogenic T cell epitopes. *Genome Medicine* 12(1):70 DOI 10.1186/s13073-020-00767-w.
- Qin L, Xiong B, Luo C, Guo ZM, Hao P, Su J, Nan P, Feng Y, Shi YX, Yu XJ, Luo XM, Chen KX, Shen X, Shen JH, Zou JP, Zhao GP, Shi TL, He WZ, Zhong Y, Jiang HL, Li YX. 2003. Identification of probable genomic packaging signal sequence from SARS-CoV genome by bioinformatics analysis. *Acta Pharmacologica Sinica* 24:489–496.
- Qiu LW, Tang HW, Wang YD, Liao JE, Hao W, Wen K, He XM, Che XY. 2005. Development and application of triple antibodies-based sandwich ELISA for detecting nucleocapsid protein of SARS-associated coronavirus. *Zhonghua Liu Xing Bing Xue Za Zhi* 26:277–281.
- Robson B. 2020. Computers and viral diseases. Preliminary bioinformatics studies on the design of a synthetic vaccine and a preventative peptidomimetic antagonist against the SARS-CoV-2 (2019-nCoV, COVID-19) coronavirus. *Computers in Biology and Medicine* 119(1):103670 DOI 10.1016/j.combiomed.2020.103670.

- Rothe C, Schunk M, Sothmann P, Bretzel G, Froeschl G, Wallrauch C, Zimmer T, Thiel V, Janke C, Guggemos W, Seilmaier M, Drosten C, Vollmar P, Zwirgmaier K, Zange S, Wolfel R, Hoelscher M. 2020.** Transmission of 2019-nCoV infection from an asymptomatic contact in Germany. *New England Journal of Medicine* **382(10)**:970–971  
DOI [10.1056/NEJMc2001468](https://doi.org/10.1056/NEJMc2001468).
- Saha R, Ghosh P, Burra V. 2021.** Designing a next generation multi-epitope based peptide vaccine candidate against SARS-CoV-2 using computational approaches. *3 Biotech* **11(2)**:47  
DOI [10.1007/s13205-020-02574-x](https://doi.org/10.1007/s13205-020-02574-x).
- Saha S, Raghava GP. 2006.** Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. *Proteins-Structure Function and Bioinformatics* **65(1)**:40–48  
DOI [10.1002/prot.21078](https://doi.org/10.1002/prot.21078).
- Sanchez CM, Izeta A, Sanchez-Morgado JM, Alonso S, Sola I, Balasch M, Plana-Duran J, Enjuanes L. 1999.** Targeted recombination demonstrates that the spike gene of transmissible gastroenteritis coronavirus is a determinant of its enteric tropism and virulence. *Journal of Virology* **73(9)**:7607–7618 DOI [10.1128/JVI.73.9.7607-7618.1999](https://doi.org/10.1128/JVI.73.9.7607-7618.1999).
- Shajahan A, Supekar NT, Gleinich AS, Azadi P. 2020.** Deducing the N- and O- glycosylation profile of the spike protein of novel coronavirus SARS-CoV-2. *Glycobiology* **30(12)**:981–988  
DOI [10.1093/glycob/cwaa042](https://doi.org/10.1093/glycob/cwaa042).
- Shaman J, Galanti M. 2020.** Will SARS-CoV-2 become endemic? *Science* **370(6516)**:527–529  
DOI [10.1126/science.abe5960](https://doi.org/10.1126/science.abe5960).
- Shang J, Ye G, Shi K, Wan Y, Luo C, Aihara H, Geng Q, Auerbach A, Li F. 2020.** Structural basis of receptor recognition by SARS-CoV-2. *Nature* **581(7807)**:221–224  
DOI [10.1038/s41586-020-2179-y](https://doi.org/10.1038/s41586-020-2179-y).
- Shen X, Xue JH, Yu CY, Luo HB, Qin L, Yu XJ, Chen J, Chen LL, Xiong B, Yue LD, Cai JH, Shen JH, Luo XM, Chen KX, Shi TL, Li YX, Hu GX, Jiang HL. 2003.** Small envelope protein E of SARS: cloning, expression, purification, CD determination, and bioinformatics analysis. *Acta Pharmacologica Sinica* **24**:505–511.
- Shulman Z, Gitlin A, Weinstein J, Lainez B, Esplugues E, Flavell R, Craft J, Nussenzweig M. 2014.** Dynamic signaling by T follicular helper cells during germinal center B cell selection. *Science* **345(6200)**:1058–1062 DOI [10.1126/science.1257861](https://doi.org/10.1126/science.1257861).
- Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Soding J, Thompson JD, Higgins DG. 2011.** Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology* **7(1)**:539  
DOI [10.1038/msb.2011.75](https://doi.org/10.1038/msb.2011.75).
- Sohail M, Ahmed S, Quadeer A, McKay M. 2021.** In silico T cell epitope identification for SARS-CoV-2: progress and perspectives. *Advanced Drug Delivery Reviews* **171**:29–47  
DOI [10.1016/j.addr.2021.01.007](https://doi.org/10.1016/j.addr.2021.01.007).
- Su S, Wong G, Shi W, Liu J, Lai ACK, Zhou J, Liu W, Bi Y, Gao GF. 2016.** Epidemiology, genetic recombination, and pathogenesis of coronaviruses. *Trends in Microbiology* **24(6)**:490–502  
DOI [10.1016/j.tim.2016.03.003](https://doi.org/10.1016/j.tim.2016.03.003).
- Sunwoo HH, Palaniyappan A, Ganguly A, Bhatnagar PK, Das D, El-Kadi AO, Suresh MR. 2013.** Quantitative and sensitive detection of the SARS-CoV spike protein using bispecific monoclonal antibody-based enzyme-linked immunoassay. *Journal of Virological Methods* **187(1)**:72–78  
DOI [10.1016/j.jviromet.2012.09.006](https://doi.org/10.1016/j.jviromet.2012.09.006).
- Tan YJ, Goh PY, Fielding BC, Shen S, Chou CF, Fu JL, Leong HN, Leo YS, Ooi EE, Ling AE, Lim SG, Hong W. 2004.** Profiles of antibody responses against severe acute respiratory syndrome coronavirus recombinant proteins and their potential use as diagnostic markers.

*Clinical Diagnostic Laboratory Immunology* **11**(2):362–371

DOI [10.1128/CDLL.11.2.362-371.2004](https://doi.org/10.1128/CDLL.11.2.362-371.2004).

- Tang MS, Hock KG, Logsdon NM, Hayes JE, Gronowski AM, Anderson NW, Farnsworth CW. 2020.** Clinical performance of two SARS-CoV-2 serologic assays. *Clinical Chemistry* **66**(8):1055–1062 DOI [10.1093/clinchem/hvaa120](https://doi.org/10.1093/clinchem/hvaa120).
- Tegally H, Wilkinson E, Giovanetti M, Iranzadeh A, Fonseca V, Giandhari J, Doolabh D, Pillay S, San EJ, Msomi N, Mlisana K, von Gottberg A, Walaza S, Allam M, Ismail A, Mohale T, Glass AJ, Engelbrecht S, Van Zyl G, Preiser W, Petruccione F, Sigal A, Hardie D, Marais G, Hsiao M, Korsman S, Davies M-A, Tyers L, Mudau I, York D, Maslo C, Goedhals D, Abrahams S, Laguda-Akingba O, Alisoltani-Dehkordi A, Godzik A, Wibmer CK, Sewell BT, Lourenço J, Alcantara LCJ, Pond SLK, Weaver S, Martin D, Lessells RJ, Bhiman JN, Williamson C, De Oliveira T. 2020.** Emergence and rapid spread of a new severe acute respiratory syndrome-related coronavirus 2 (SARS-CoV-2) lineage with multiple spike mutations in South Africa. *MedRxiv* DOI [10.1101/2020.12.21.20248640](https://doi.org/10.1101/2020.12.21.20248640).
- Thabet L, Mhalla S, Naija H, Jaoua MA, Hannachi N, Fki-Berrajah L, Toumi A, Karray-Hakim H. 2020.** SARS-CoV-2 infection virological diagnosis. *Tunis Med* **98**:304–308.
- Trolle T, Metushi IG, Greenbaum JA, Kim Y, Sidney J, Lund O, Sette A, Peters B, Nielsen M. 2015.** Automated benchmarking of peptide-MHC class I binding predictions. *Bioinformatics* **31**(13):2174–2181 DOI [10.1093/bioinformatics/btv123](https://doi.org/10.1093/bioinformatics/btv123).
- Vita R, Mahajan S, Overton JA, Dhanda SK, Martini S, Cantrell JR, Wheeler DK, Sette A, Peters B. 2019.** The immune epitope database (IEDB): 2018 update. *Nucleic Acids Research* **47**(D1):D339–D343 DOI [10.1093/nar/gky1006](https://doi.org/10.1093/nar/gky1006).
- Walls AC, Xiong X, Park YJ, Tortorici MA, Snijder J, Quispe J, Cameroni E, Gopal R, Dai M, Lanzavecchia A, Zambon M, Rey FA, Corti D, Veesler D. 2019.** Unexpected receptor functional mimicry elucidates activation of coronavirus fusion. *Cell* **176**(5):1026–1039.e1015 DOI [10.1016/j.cell.2018.12.028](https://doi.org/10.1016/j.cell.2018.12.028).
- Wang P, Sidney J, Kim Y, Sette A, Lund O, Nielsen M, Peters B. 2010.** Peptide binding predictions for HLA DR, DP and DQ molecules. *BMC Bioinformatics* **11**(1):568 DOI [10.1186/1471-2105-11-568](https://doi.org/10.1186/1471-2105-11-568).
- Wang R, Hozumi Y, Yin C, Wei G-W. 2020a.** Mutations on COVID-19 diagnostic targets. *Genomics* **112**(6):5204–5213 DOI [10.1016/j.ygeno.2020.09.028](https://doi.org/10.1016/j.ygeno.2020.09.028).
- Wang R, Hozumi Y, Yin C, Wei GW. 2020b.** Decoding SARS-CoV-2 transmission and evolution and ramifications for COVID-19 diagnosis, vaccine, and medicine. *Journal of Chemical Information and Modeling* **60**(12):5853–5865 DOI [10.1021/acs.jcim.0c00501](https://doi.org/10.1021/acs.jcim.0c00501).
- Wilkins MR, Gasteiger E, Bairoch A, Sanchez JC, Williams KL, Appel RD, Hochstrasser DF. 1999.** Protein identification and analysis tools in the ExpASY server. *Methods in Molecular Medicine* **112**:531–552 DOI [10.1385/1-59259-584-7:531](https://doi.org/10.1385/1-59259-584-7:531).
- Woo PC, Lau SK, Wong BH, Tsoi HW, Fung AM, Kao RY, Chan KH, Peiris JS, Yuen KY. 2005.** Differential sensitivities of severe acute respiratory syndrome (SARS) coronavirus spike polypeptide enzyme-linked immunosorbent assay (ELISA) and SARS coronavirus nucleocapsid protein ELISA for serodiagnosis of SARS coronavirus pneumonia. *Journal of Clinical Microbiology* **43**(7):3054–3058 DOI [10.1128/JCM.43.7.3054-3058.2005](https://doi.org/10.1128/JCM.43.7.3054-3058.2005).
- Yan Y, Chang L, Wang L. 2020.** Laboratory testing of SARS-CoV, MERS-CoV, and SARS-CoV-2 (2019-nCoV): current status, challenges, and countermeasures. *Reviews in Medical Virology* **30**(3):e2106 DOI [10.1002/rmv.2106](https://doi.org/10.1002/rmv.2106).
- Yuan S, Chan HCS, Filipek S, Vogel H. 2016.** PyMOL and Inkscape bridge the data and the data visualization. *Structure* **24**(12):2041–2042 DOI [10.1016/j.str.2016.11.012](https://doi.org/10.1016/j.str.2016.11.012).

- Zhan SH, Deverman BE, Chan YA. 2020.** SARS-CoV-2 is well adapted for humans. What does this mean for re-emergence? *BioRxiv* **30(4)**:343 DOI [10.1101/2020.05.01.073262](https://doi.org/10.1101/2020.05.01.073262).
- Zhang L, Udaka K, Mamitsuka H, Zhu S. 2012.** Toward more accurate pan-specific MHC-peptide binding prediction: a review of current methods and tools. *Briefings in Bioinformatics* **13(3)**:350–364 DOI [10.1093/bib/bbr060](https://doi.org/10.1093/bib/bbr060).
- Zheng K, Bantog C, Bayer R. 2011.** The impact of glycosylation on monoclonal antibody conformation and stability. *MAbs* **3(6)**:568–576 DOI [10.4161/mabs.3.6.17922](https://doi.org/10.4161/mabs.3.6.17922).