Methods to Identify Immunogenic Peptides in SARS-CoV-2 Spike and Protective Monoclonal Antibodies in COVID-19 Patients

Lili Li, Meiling Gao, Jie Li, Shulong Zu, Yanan Wang, Chunfeng Chen, Dingyi Wan, Jing Duan, Roghiyh Aliyari, Jingfeng Wang, Jicai Zhang, Yujie Jin, Weijin Huang, Xiaoxia Jin, Minxin Shi, Youchun Wang, Cheng-Feng Qin, Heng Yang,* and Genhong Cheng*

1. Introduction

The global pandemic of a novel beta-coronavirus, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has caused great health and economic impacts. Infection with SARS-CoV-2 results in coronavirus disease 2019 (COVID-19), from which over 3.2 million people have died.[1–3] To date, many potential treatment approaches have been explored based on the progress of SARS-CoV-2 research. Among them, neutralizing antibodies that inhibit SARS-CoV-2 fusion/entry are promising.[4–8]

In response to SARS-CoV-2 infection, patients generate polyclonal antibodies that recognize numerous epitopes across the viral proteome. This pool of polyclonal antibodies (including IgM, IgG, etc.) provides humoral immunity against...
SARS-CoV-2. Sera from recovered COVID-19 patients can neutralize SARS-CoV-2 in an in vitro assay and have been successfully used to treat other COVID-19 patients.

The surface spike glycoprotein of SARS-CoV-2 is critical for virus infection through engaging the host receptors and mediating virus-host membrane fusion. The receptor-binding domain (RBD) is located in SARS-CoV-2 S1 domain, which binds to the cellular receptor angiotensin-converting enzyme 2 (ACE2) through the interaction with its receptor-binding motif (RBM).

Interestingly, the RBD domain is also thought to be the major antigen of SARS-CoV-2 and is the critical target for neutralizing antibodies. Although SARS-CoV-2 and SARS-CoV share high levels of sequence homology across their respective proteomes, they are more divergent in the S protein sequences, especially in the RBM, which shares only 59% homology. This suggests the possibility that COVID-19 patients may develop SARS-CoV-2 specific monoclonal antibodies, including neutralizing antibodies against the RBM domain of SARS-CoV-2 with limited cross-reactivity to SARS-CoV.

A number of neutralizing antibodies against SARS-CoV-2 have been and are being developed as therapeutic agents, which can be used individually such as hamlanivimab or as a cocktail such as REGN-COV2 in the treatment of COVID-19. Besides their vaccination and therapeutic applications, antigens and antibodies are also great tools for COVID-19 diagnosis. While symptoms and computed tomography (CT) scans determine disease status and PCR-based swab testing determine the levels of viral RNA, blood antibody testing measures the levels of SARS-CoV-2 IgM and IgG antibodies in COVID-19 patients. Although it may take over a week or two for the antibody response to be present, blood antibody testing may provide more reliable results for the diagnosis of recovered or asymptomatic COVID-19 patients. As SARS-CoV-2 shares significant sequence homology with SARS-CoV, it is expected that they also share certain epitopes, which could lead to the generation of cross-reactive antibodies. Though cross-reactive antibodies can be cross-protective, they can also confound diagnostic antibody testing.

The antibody protective effectiveness and the disease outcomes of COVID-19 patients are determined both by the serum levels of antibodies and by the epitopes they recognize. However, it is not clear how many immunogenic epitopes exist among SARS-CoV-2 viral proteins and more importantly which epitopes can generate protective antibodies against SARS-CoV-2 infection. In this study, we have built a phage display library containing overlapping peptides across the SARS-CoV-2 spike protein and identified numerous immunogenic peptides, including three immunogenic peptides within the RBM, with strong antibody response in COVID-19 patients. To directly identify human monoclonal antibodies from patients, we have also generated another phage display library expressing pairs of antibody heavy and light chains from peripheral blood mononuclear cells of COVID-19 patients and identified immune protective antibodies against SARS-CoV-2 infection.

2. Results

In response to SARS-CoV-2 infection, patients generate polyclonal antibodies that recognize numerous epitopes across the viral proteome. Among the surface proteins of SARS-CoV-2, spike protein is the most antigenic and the best candidate for strain-specific amino acid epitopes. To confirm the generation of antibodies against spike protein in COVID-19 patients, we measured the IgG and IgM levels against spike protein in COVID-19 patients (38 cases) and healthy (225 cases) control people’s sera. The results showed higher antispike antibody levels in COVID-19 patients, while no obviously antibodies against SARS-CoV-2 were detected in healthy people sera (Figure 1a). To systematically identify immunogenic epitopes within the spike protein of SARS-CoV-2, we constructed spike protein phage display library expressing 25-amino acid peptides which cover the full-length spike protein (Figure 1b). This phage library was incubated with sera from COVID-19 patients and healthy individuals (Table S1, Supporting Information) and the antibody binding phages were then enriched by protein A/G beads for subsequent PCR amplification and sequencing. Through this phage display analysis, we have identified about 15 immunogenic peptides that can detect antibodies in multiple COVID-19 patients. Among them, we found three peptides located in the RBM domain had obviously antibody reactivity in all the tested COVID-19 patients (Figure 1c).

We further used Immune Epitope Database and Analysis Resource (IEDB) prediction to confirm the antigenicity of the three peptides and predict the sequences (Figure 1d). Based on the computer-guided homology modeling method, the interaction structural models of SARS-CoV-2 S protein (PDB ID: 6M0J) and ACE2 protein were constructed by SWISS-MODEL online server (Figure 1e). The structural model indicated that these three peptides were exposed on the surface of the S protein. When we compared the amino acid sequences in these regions of SARS-CoV-2 to those in the corresponding regions of SARS, MERS, and other human coronaviruses, we saw that the three peptide sequences of the SARS-CoV-2 strain differ significantly from those of related coronaviruses (Figure 1f). These results suggest that there are at least three immunoreactive epitopes in the RBM of the S protein, which may generate specific antibodies in COVID-19 patients.

To investigate the antibody profiles of different patients, we first used serum samples from healthy individuals, COVID-19 suspected patients and COVID-19 patients (Table S2, Supporting Information) to measure both the IgM and IgG levels against the three peptides and a control peptide (S25-39). As shown in Figure 2a–f, the IgG and IgM antibody levels specific for the three peptides in COVID-19 suspected patients and COVID-19 patients were significantly higher than the healthy individuals. While most of the COVID-19 patients were IgG positive, fewer of them were IgM positive. Interestingly, in COVID-19 suspected patients who were PCR- at the time of serum sample collection but had COVID-19 like symptoms and in some cases confirmatory lung CT scans, the levels of IgG and IgM antibody were also significantly higher than healthy individuals (Figure 2a–d). Our heatmap analysis has also demonstrated that the majority of patients who had COVID-19 like symptoms and/or CT scan were positive while healthy individuals were all negative with IgG and IgM against the three peptides, regardless whether they were PCR+ or PCR- for SARS-CoV-2 RNA at the time of serum sample
The results indicated that these three peptide epitopes detected IgG and IgM antibodies in both COVID-19 patients and COVID-19 suspected patients.

To further explore the possibility of using the three peptides as antigens for the diagnosis of COVID-19 patients, we compared the IgG and IgM levels of both COVID-19 patients and COVID-19 suspected patients with those of healthy individuals.

**Figure 1.** Screening and identification of B-cell epitopes in SARS-CoV-2 spike protein. a) Serum levels of IgG and IgM antibodies against SARS-CoV-2 spike protein in 38 COVID-19 patients and 225 healthy people were detected by ELISA assays. b) Overview of the peptide screening procedure. SARS-CoV-2 spike protein peptide library includes 25-mer peptides across the spike protein. Oligonucleotides were cloned into T7 bacteriophage display vector and packaged into phage particles displaying the encoded peptides on their surface. The phage library was mixed with sera from 7 COVID-19 patients and one healthy individual containing antibodies that bind to their cognate epitopes. Bounded phage were isolated by immunoprecipitation with protein A/G coated magnetic beads. Last, PCR amplification and Illumina sequencing from the DNA of the bound phage were performed to reveal the peptides targeted by the serum antibodies. c) Antibody profile of randomly chosen COVID-19 patients and healthy people. Each row represents a SARS-CoV-2 spike protein peptide, each column represents a serum sample. The color intensity of each cell indicates the log10 (reads counts+1) of peptides from SARS-CoV-2 spike protein. d) Three peptides located within the receptor-binding motif (RBM) region were significantly enriched in COVID-19 patients and the predicted B-cell epitope are listed in the table. e) Overall structure of SARS-CoV-2 RBD and ACE2. ACE2 is shown in green. The SARS-CoV-2 RBD core is shown in cyan and S431-454 in yellow, S470-486 in orange, S501-515 in purple. The Protein Data Bank (PDB) code for the SARS-CoV-2 RBD-ACE2 complex is 6M0J. f) Sequence comparison of the three peptides with human and bat SARS-like CoVs and MERS. The ELISA data in (a) are means ± SD, **** P < 0.0001, paired Student’s t test.
Both IgM and IgG levels were significantly higher in COVID-19 patients than in healthy individuals. The levels of IgM and IgG in COVID-19 suspected patients, although lower than those in COVID-19 patients, were significantly higher than those in healthy individuals (Figure 3a–d). Figure 3c,d represent the receiver operating characteristics (ROC) curves for measuring both the sensitivity and specificity of the three peptides and S25-39 control peptide in the diagnosis of both COVID-19 patients and COVID-19 suspected patients. The areas under the ROC (AUC) for S25-39, S431-454, S470-486 and S501-515 peptides IgG were 0.6432, 1, 0.994, and 0.9974 among COVID-19 patients and 0.6286, 0.981, 0.9991 and 0.982 among COVID-19 suspected patients, respectively (Figure 3c). The AUC values for S25-39, S431-454, S470-486 and S501-515 peptides IgM were 0.5428, 0.865, 0.9588 and 0.9712 among COVID-19 patients and 0.5795, 0.9805, 0.9705 and 0.981 among COVID-19 suspected patients, respectively (Figure 3d). These results suggest that the levels of antibodies against the three peptides, especially those for IgG antibodies, can be used as biomarkers for the diagnosis of COVID-19 patients.

To determine the relationship between antibody levels and clinical status of individual patients, we first monitored the SARS-CoV-2 pseudovirus neutralization ability of four COVID-19 patients' sera (Table S3, Supporting Information) along with their clinical scores, which represent the overall patient health conditions including both blood test results and clinical symptoms such as fever, cough, shortness of breath and indications of tissue injury in CT scans. All four COVID-19 patients developed severe symptoms but three out of four patients recovered, while patient #1 died. We observed a potential reverse correlation between the pseudovirus neutralization ability of four patients' sera and the clinical scores of individual patients despite different outcomes (Figure S2, Supporting Information). We further detected IgG levels against the S431-454, S470-486, and S501-515 peptides in the four patients and found that the antibody levels changed during the hospital stay. Interestingly, although the antibody levels against S470-486 peptide were the highest among the three peptides, the antibody profiles of individual peptides seemed not to match well with the patients' neutralization activity (Figure S2, Supporting Information). These results suggest that the overall neutralization activity of COVID-19 patients is likely resulted by a combination of antibodies against multiple epitopes.

The three peptides from the RBM domain of SARS-CoV-2 spike protein not only showed immunoreactivity in COVID-19 patients but also are involved in contacting with ACE2 during
SARS-CoV-2 infection (Figure 1e). To test the neutralization ability of these three peptides, we pretreated Huh7.5 cell with BSA-conjugated synthetic S431-454, S470-486, and S501-515 peptides and infected the cells with SARS-CoV-2 pseudovirus. The result showed that the three BSA-conjugated synthetic peptides remarkably blocked SARS-CoV-2 pseudovirus infection (Figure 4a). Similarity, while transfection of hACE2 plasmid into HEK293T cells strongly enhanced SARS-CoV-2 pseudovirus infection, the addition of the BSA-conjugated synthetic three peptides inhibited hACE2-mediated SARS-CoV-2 pseudovirus infection (Figure 4b). Furthermore, we found that BSA-conjugated S431-454 and S470-486 peptides significantly inhibited wild-type SARS-CoV-2 infection (Figure 4c). To further determine if antibodies raised by individual peptides have neutralization effects against SARS-CoV-2 infection, we immunized mice with the three peptides in the presence of Freund adjuvant. Interestingly, we only detected high
levels of S431-454 but not S431-454 and S501-515 specific antibodies present in the sera of immunized mice (Figure 4d).

We used the sera from S431-454, S470-486, and S501-515 immunized mice to test neutralization activity against SARS-CoV-2 pseudovirus. As shown in Figure 4e,f, sera from S470-486 immunized mice but not S431-454, S501-515, and the unimmunized mice strongly inhibited the infection of HuH7.5 cells and hACE2 transfected HEK293T cells with SARS-CoV-2 pseudovirus. When we further tested the viral neutralization activity of S470-486 immunized mice sera with SARS-CoV-2 authentic virus, the result showed that sera from S470-486 immunized mice can also inhibit the infection of SARS-CoV-2 authentic virus (Figure 4g). These results suggest the S470-486 peptide contains an immunogenic epitope, which can raise SARS-CoV-2 neutralization antibodies upon immunization in mice.

To identify human monoclonal antibodies directly from COVID-19 patients that can specifically bind to S470-486 peptide, we have constructed ScFv phage display library expressing pairs of Ig light and heavy chain variable regions isolated from the cDNA pool of the 15 COVID-19 patient PBMC samples. After
three rounds of enrichment from the library containing $8.7 \times 10^9$ ScFv phages for their ability to bind to biotin-labeled S470-486 peptide, we have obtained two clones (R3P1-F8 and R3P2-A9-3), which were reconstructed into the expression vector containing the human IgG1 backbone (Figure 5a). We further confirmed the binding ability of full-length R3P1-F8 and R3P1-A9-3 antibodies to the S470-486 peptide by ELISA assay (Figure 5b,c). More importantly, R3P1-F8 IgG1 antibodies effectively blocked the infection of SARS-CoV-2 pseudovirus and wild-type virus (Figure 5d,e). These results have therefore demonstrated that the phage display method was an effective way to identify human clonal antibodies directly from the COVID-19 patients and the resulting monoclonal antibodies against S470-486 peptide were able to neutralize SARS-CoV-2 infection.

3. Discussion

The SARS-CoV-2 infection and its associated COVID-19 disease have been declared as pandemic by World Health Organization (WHO). The COVID-19 pandemic rapidly spreads across the world in a few months and is an emerging threat to global public health. To face such challenge, we have developed a comprehensive approach by taking the advantages of both peptide-based and antibody-based phage display libraries to screen for immunogenic epitopes and their corresponding human monoclonal antibodies. We have also used peptide-based ELISA to measure antibody levels in COVID-19 patients and immunized mice as well as viral infectivity assays to characterize the functional activities of these epitopes and antibodies.

VirScan and proteome peptide microarray have been developed to systematically analyze epitope of antiviral antibodies in COVID-19 patients’ sera.[31,32] Shrock et al. used VirScan to screen sera from 232 COVID-19 patients and mapped 823 distinct epitopes across the entire SARS-CoV-2 proteome and found 10 of which are likely targets of neutralizing antibodies.[31] We used a similar method to construct a peptide-based phage display library expressing 25-amino acid peptides, which cover the full-length spike protein and identified about 15 immunogenic peptides that can detect antibodies in multiple COVID-19 patients. Wang et al. developed a proteome peptide microarray and screened sera from 10 COVID-19 patients and identified a peptide epitope for neutralizing antibodies within the SARS-CoV-2 spike RBM’s interaction interface with ACE2.[32] Due to different COVID-19 patients’ sera used and different peptide microarray construction strategy, diverse epitope peptides were identified by different research groups and novel epitope is being explored.
We have further characterized the three immunogenic peptides within the RBM of SARS-CoV-2 spike protein through combing peptide phage display of SARS-CoV-2 spike protein with antigen epitope bioinformatics prediction. Based on the structural prediction, the three peptides in the RBM are involved in interaction with the ACE2 receptor (Figure 1e). Antibodies binding to these peptides will likely neutralize viral infection through directly competing with ACE2. Therefore, we may potentially use the three peptides as diagnosis tools to measure functional antibody levels in COVID-19 patients against SARS-CoV-2 infection.

Although PCR based assays to detect viral RNA in the patients’ oral or nasal swabs are the standard method for COVID-19 diagnosis,[33] their efficacy is limited to a narrow window of active infection period and confound by low detectability, as most of SARS-CoV-2 infection happened in the deep lung tissues. Therefore, it is urgent to develop specific, sensitive, and reliable alternative methods for COVID-19 diagnosis as well as surveillance of asymptomatic or recovered patients. It is also equally important to develop tools to measure protective immunity in recovered COVID-19 patients or vaccinated individuals. Although detecting antibody levels against SARS-CoV-2 have been used for COVID-19 diagnosis. Most of the antigens in various diagnostic kits were fragments from spike (S), nucleocapsid (N), or membrane (M) proteins, which may detect cross-reactive antibodies in patients previously infected with other coronaviruses.[34–37] Furthermore, detecting antibodies against these fragments does not necessarily tell the protective immunity in COVID-19 patients or vaccinated individuals. Through phage display analysis and structural-based prediction, we have found three peptides within the RBM domain of SARS-CoV-2 spike protein with low homology to other coronaviruses and tested their ability to detect antibodies in COVID-19 patients (Figure 1b–d). Interestingly, all the three peptides detected antibodies in both COVID-19 patients or suspected patients and the antibody levels fluctuated with the COVID-19 disease progress (Figure S2, Supporting Information). As the three peptides are in the regions involved in direct interaction with the entry receptor ACE2, they may be useful in indicating protective immunity in COVID-19 patients. However, the current study remains limited by a small number of samples without samples from asymptomatic infected patients or SARS-CoV-2 vaccinated individuals. Future studies will be needed to determine the efficiency of these antigens to detect recovered COVID-19 patients and vaccinated individuals who may gain protective immunity against SARS-CoV-2 infection.

In addition to measuring the antibody levels against these three peptides in the RBM in COVID-19 patients, we have further determined their abilities to inhibit SARS-CoV-2 infection and their immunogenicity in mice. Our results showed that these three peptides in cell permeable forms can all effectively inhibit the infection of infection pseudo- and real SARS-CoV-2.

Interestingly, when we immunized Balb/c mice with the three peptides to stimulate the production of corresponding antibodies, we only detected antibodies binding to S470-486 peptide in the sera of S470-486 peptide immunized mice (Figure 4d). This indicated that S470-486 peptide was a stronger immunogenic epitope than S431-454 and S501-515. Similarly, the levels of antibodies binding to S470-486 peptides were the highest among the three peptides in the sera of four COVID-19 patients with different disease progresses (Figure S2, Supporting Information). S431-454 and S501-515 peptides did not stimulate the production of neutralizing antibody (Figure 4d), which may result from the experiment protocols, the difference of immune system between mice and human, or the difference between SARS-CoV-2 infection and immunogenic peptide.

Traditional methods to obtain therapeutic human monoclonal antibodies require immunizing mice with the antigen of interest, generating hybridoma cell lines, screening for mouse monoclonal antibodies against the antigen of interest and humanizing monoclonal antibodies toward human monoclonal antibodies. The entire process may take several years requiring significant efforts and resources. In contrast, the method we used in the current studies can obtain fully human monoclonal antibodies directly from patients within about three weeks. Our method of identifying human antibodies against the RBM of SARS-CoV-2 involved in isolation of PBMC from COVID-19 patients, constructing the ScFv phage display library expressing pairs of Ig light and heavy chain variable regions, screening for ScFv phages that can bind to biotin-labeled peptide, and reconstructing antigen-specific ScFv into the expression vector containing the human IgG1 backbone (Figure 5a). We have demonstrated that the resulting monoclonal antibodies not only strongly bound to the specific antigen but also effectively blocked the infection of SARS-CoV-2 pseudovirus and wild-type virus (Figure 5b–e). Of course, more research work is needed about the efficiency, safety, and potential of clinical application of these isolated monoclonal antibodies.

In the past two decades, we have encountered unpredictable emerging and highly pathogenic viruses, including SARS, Ebola, Zika, and SARS-CoV-2. It is critical to explore novel tools that can quickly identify functional epitopes within the viral proteomes and isolate monoclonal antibodies in order to develop reagents for new diagnosis, vaccine, and therapeutic antibodies. Our studies have outlined a comprehensive approach to use peptide-based phage display library for immunogenic epitope screening, various molecular and cellular assays for functional validation and ScFv phage display library for isolating monoclonal antibodies from patients. The approach we used for SARS-CoV-2 should also be useful for other viruses.

4. Experimental Section

**Virus, Cells, Plasmids, and Reagents:** SARS-CoV-2 virus (GenBank accession number Txd2697049) was used in this study. SARS-CoV-2 pseudovirus with Luciferase coding sequence was constructed as described.[38] Huh7.5, HEK293T, and Vero cell lines were purchased from American Type Culture Collection and cultured in Dulbecco’s modified Eagle’s medium (DMEM) (37 °C, 5% CO₂) supplemented with 10% fetal bovine serum (FBS), 100 U mL⁻¹ penicillin, and 50 μg mL⁻¹ streptomycin. hACE2 and hTMPRSS2 expression plasmids were purchased from GenScript (Nanjing, China). MES monohydrate was purchased from Sigma, 1-(3-dimethylamino propyl)-3-ethyl carbodiimide hydrochloride (EDC) and 3,3′,5,5′-tetramethylbenzidine (TMB) was purchased from ThermoFisher Scientific.

**Clinical Investigations and Data Analysis:** Clinical investigations in patients with SARS-CoV-2 infection and healthy volunteers were approved by the Ethics Committee of Taihe hospital (2020KS043). Sera...
originating from human donors were collected after the patients were admitted to the hospital with obvious SARS-CoV-2 infection symptoms like fever, with or without recorded temperature, radiographic evidence of pneumonia, low or normal white-cell count or low lymphocyte count and positive SARS-CoV-2 nucleic acid testing. All sera were collected and used under a protocol approved by the local governing human research protection committee of Taihe hospital. Donors are divided into three groups according to SARS-CoV-2 PCR results and clinical symptoms: COVID-19 patients were diagnosed because of PCR+ and infection symptoms \((n = 43)\), COVID-19 suspected patients were PCR- at the time we obtained sera, but with SARS-CoV-2 infection symptoms \((n = 117)\), healthy people were PCR- and without any SARS-CoV-2 infection symptoms \((n = 38)\). A modified Acute Physiology and Chronic Health Evaluation (APACHE II) scoring system was used to evaluate the disease severity of four patients during the hospitalization. Briefly, APACHE II is a severity of disease classification system based on 12 physiologic measurements (body temperature, mean arterial pressure, heart rate, respiratory rate, alveolar-arterial oxygen \((A-a)\) gradient; if the fractional inspired oxygen concentration is \(\geq 0.5\), arterial oxygen tension \((PaO_2)\); if the fractional inspired oxygen concentration is \(<0.5\), serum bicarbonate \((HCO_3^-)\); if there is no arterial blood gas analysis, arterial \(pH\), serum potassium, creatinine, hematocrit, white blood cell count, Glasgow Coma Scale score), age of the subjects and comorbid conditions. Physiologic signs and laboratory values used to calculate the APACHE II score were obtained using the worst physiologic variable within the 24 h period. A higher score is associated with more severe disease and a higher risk of hospital death.

**Screening for SARS-CoV-2 Immunogenic Peptides by Phage Display Library:** For the coronavirus spike protein library construction, the DNA sequence of SARS-CoV-2 spike gene was first extracted in National Center for Biotechnology Information (https://www.ncbi.nlm.nih.gov/). Then 25-amino acid peptide sequences covered the full-length spike protein were DNA codons optimized for expression, and cloned into the N-terminus of PIIL protein bacteriophage M13. Then, the phage library was mixed with sera from 7 COVID-19 patients and one healthy individual to screen for immunogenic peptides. After incubation at room temperature \((RT)\) for 1 h and left the mixture at \(RT\) for another 1 h, bounded phage was isolated by immunoprecipitation with protein A/G coated magnetic beads (Invitrogen). After incubation on a rotator at \(RT\) for 20 min, the beads were washed with PBS \((1×\) PBS supplemented with 0.05% Tween-20) for eight times and PBS for two times. Then, 300 \(\mu\)L 1 mg/mL trypsin \((\text{Trypsin; PBS = 1.5})\) was added and incubated in the tube at \(37\ °C\) for 10 min. Repeated the digestion step for two times. Then, the number of phages bound to beads was evaluated by the phage-plaque assay. The phages were amplified in TG1 for next round of biopanning. After three rounds of biopanning, PCR amplification and Illumina sequencing from the DNA of the bound phage were performed to reveal the peptides targeted by the serum antibodies. The primers used for the first round of PCR amplification were forward: AACACTTCTTTCCATCAACAGCCTCTTCGATCTGATTGCAGTGGCAGCTGCTTCCGATCTACGGGTATGCGCCATGGTGATGGTGATGGTG. The second round of PCR was performed with products from the first-round PCR as templates. The primers used for the second round of PCR amplification were forward: AATGATACGGCGACCACCGAGCTCTTCCGATCTGATTGCAGTGGCAGCTGCTTCCGATCTACGGGTATGCGCCATGGTGATGGTGATGGTG and reverse: GACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCCGGCAGGTAGTGGTGATGGTG.

**Viral Sequence Analysis:** Spike protein sequences of representative human CoVs, but SARS-like CoVs, MERS, and SARS-CoV-2 were obtained from NCBI GenBank (The corresponding accession numbers are shown in Figure 1F). The multiple sequence alignment was conducted by Clustal Omega method using the MegAlign Pro software in DNASTAR Lasergene package with default parameters.

**Enzyme-Linked Immunosorbent Assay:** Enzyme-Linked Immunosorbent Assay (ELISA) was performed with modifications. First, synthesized the peptides from GenScript (Nanjing, China) and diluted in MES buffer \((0.1 \, \text{m}, \text{PH = 6.0})\). The plates were coated with 10 \(\mu\)L EDC \((10 \, \text{mg} \, \text{mL}^{-1})\) and 50 \(\mu\)L peptide \((4 \, \mu\text{g} \, \text{mL}^{-1})\) per well and incubated at 4 °C for overnight. After washing three times, the coated plates were blocked with 100 \(\mu\)L per well blocking solution and incubated at 37 °C for 1 h. Diluted sera \((1:500)\) at 100 \(\mu\)L per well were added to the plates and incubated at 37 °C for 2 h. After washing three times. Diluted HRP-goat anti-human IgG secondary antibody at 100 \(\mu\)L per well was added and incubated for 1 h. Washing five times. TMB at 100 \(\mu\)L per well was added into the plates and incubated for 5–10 min with light protection and the reaction was terminated by the addition of 50 \(\mu\)L per well 2 M \(\text{H}_2\text{SO}_4\). The absorbance of each well at 450 nm was measured with SpectraMax i3 ( Molecular Devices) plate reader.

The antibodies against spike protein in COVID-19 patients and healthy individuals’ sera were detected by ELISA with full length spike protein (Tarcine BioMed Inc).

**Net OD450 Calculation**: Net OD450 values were used in statistical analysis of ELISA data and in sensitivity and specificity calculations for ROC curves. The net OD450 is equal to the OD450 in the experimental group minus the background OD450 measured in control group in which sera were added to wells without peptide.

**Receiver Operating Characteristic Curve Analysis**: The receiver operating characteristic (ROC) curves, their corresponding areas under the curve (AUC), and the specificities and sensitivities of IgG and IgM were established by testing 117 COVID-19 suspected patients and 43 COVID-19 patients. ROC, AUC, sensitivities, and specificities were calculated using the predicted values estimated by supervised machine learning random forest (RF) algorithm models. The classification RF algorithm is a supervised machine learning method that is used to predict the class of an observation based on the knowledge that is previously acquired from the training data. By seeing many observations, with the value of each predictor variable and the classification results, the algorithm is trained and is able to classify a previously unseen observation. The AUC and cutoff value were calculated using MedCalc statistical software.

**Mice Challenge and Immunization**: Balb/c mice were purchased from Beijing Vital River Laboratory Animal Technology Co., Ltd. All animals were bred in the core animal facility and all the animal experiments were approved by the Institutional Animal Care and Use Committee of Suzhou Institute of Systems Medicine (ISM-IACUC-0019-R). Each mouse was immunized with 50 \(\mu\)g BSA-conjugated peptide mixed with 50 \(\mu\)L Freunds adjuvant. Eight days after immunization, 100 \(\mu\)L of orbital blood was collected with glass capillary and serum was isolated and stored at −80 °C.

**Virus Neutralization Assay**: The SARS-CoV-2 pseudovirus neutralization ability of sera and R3P1-F8 and R3P2-A9-3 antibodies was determined in Huh7.5 cells via Luciferase reporter assay. Briefly, serial diluted sera or antibodies were mixed with SARS-CoV-2 pseudovirus and incubated at 37 °C for 2 h. The above mixture was then added to Huh7.5 cells and incubated for 24 h. Luciferase activity was measured by Luciferase reporter assay as described previously with the average number in the control group set as 0, representing no neutralization.

**Peptide Inhibition Assay**: Huh7.5 cells were seeded in 48 well plates and cultured overnight. BSA-conjugated 5431-454, 5470-486, and 5501-515 peptides were diluted as indicated and added to Huh7.5 cells. Huh7.5 cells were cultured for 2 h and then infected with SARS-CoV-2
pseudovirus. 24 h post infection, luciferase activity was measured by luciferase reporter assay and the pseudovirus inhibition ratio was calculated.

**COVID-19 ScFv Phage Display Library Construction and Biopanning:** To construct COVID-19 ScFv phage display library, 15 COVID-19 patients’ PBMC samples were collected and used to do RNA extraction and cDNA reverse transcription. The VH and VL gene fragments were PCR amplified using the mixed 15 COVID-19 patients’ cDNA as template. The amplified VL gene fragments were cut with restriction enzymes Nhel and Not1, agarose gel-purified, and ligated into the vector pATA-ScFv-2 which including the M13 Gene III, ScFv linker (GGGCGGGGSSCSCGGGCGG), multiple restriction sites, Lac promoter, Lac operator and pel B signal peptide. The ligated DNA mixture was electroporated into Escherichia coli TG1 cells. The transformed TG1 clones were collected and the vector with the VL gene fragments was extracted. The amplified VH gene fragments were cut with restriction enzymes SfiI and Xhol, agarose gel-purified and ligated into the vector with the VL gene fragments. The ligated DNA mixture was electroporated into E. coli TG1 cells. Ninety-six transformed TG1 clones were picked to do the colony PCR and sequencing validation. The rest of transformed TG1 clones were collected and used to amplify the library phages. The final titer of COVID-19-pATA-ScFv phage display library capacity is 8.7 × 10^8 CFU.

**Library Phages were rescued and amplified by expanding culture transformed TG1 cells with M13KO7 helper phages.** Specific phages against the S470-486 from the COVID-19-pATA-ScFv phage display library were affinity-enriched by four rounds biopanning which cross-used biotin against the S470-486 from the COVID-19-pATA-ScFv phage display library.

**Quantitative Reverse Transcription PCR:** Vero cells were treated with either dH2O (vehicle control) or 1 µg mL^-1^ of BSA-conjugated S431-454, S470-486, and S501-515 peptides for 1 h and then infected with 100 µL TCD50 SARS-CoV-2. 24 h postinfection, the SARS-CoV-2 RNA copies in culture supernatant were measured by quantitative reverse transcription-PCR (qRT-PCR).

**Statistics:** Statistical analyses were performed with GraphPad Prism 7 software and R Studio version 3.6.3. The continuous variables were presented as mean ± SD. Data with normal distribution were analyzed by one-way ANOVA or paired two-tailed Student’s t tests, and P values were indicated by ns, not significant, *P < 0.05, **P < 0.01, ***P < 0.001, and ****P < 0.0001.

**Ethics Statement:** In this study, 160 COVID-19 patients and 225 healthy individuals were enrolled. Venous blood sampling was performed to obtain sera. This study was approved by Ethics Committee of Taihe hospital (2020K5043). All subjects signed an informed consent prior to enrolment. The animal study was reviewed and approved by The Ethics Committee of Suzhou Institute of Systems Medicine.

**Conflict of Interest**

Y.W. and C.C. were employed by the company Suzhou Func Biotech Inc. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Author Contributions**

L.L., M.G., J.L., S.Z., and Y.W. contributed equally to this work. G.C., H.Y., and L.L. jointly designed this study. Y.W., M.G., J.L., J.Z., Y.J., and M.S. provided the SARS-CoV-2 pseudovirus. D.W. and J.D. performed phage display library screening and ScFv purification assay. R.A., and S.Z. performed the virus neutralization assay. L.L. and J.W. collected patient serum samples and clinical data analysis. H.Y., and L.L. jointly designed this study. Y.W., M.G., J.L., S.Z., and Y.W. contributed equally to this work. G.C., H.Y., and L.L. jointly designed this study. Y.W., M.G., J.L., J.Z., Y.J., and C.C. performed ELISA and mice challenge and immunization assay. L.L. and J.W. collected patient serum samples and clinical data analysis. D.W. and J.D. performed phage display library screening and ScFv purification assay. R.A., and S.Z. performed the virus neutralization assay. W.H., Y.W., X.J., C.-F.Q., and M.S. provided the SARS-CoV-2 pseudovirus.

**Data Availability Statement**

The data that supports the findings of this study are available in the supplementary material of this article.

**Keywords**

coronavirus disease 2019, immunogenic peptides, neutralizing antibodies, severe acute respiratory syndrome coronavirus 2, spike proteins

Received: January 19, 2021
Revised: May 19, 2021
Published online: June 19, 2021
[39] Q. Li, X. Guan, P. Wu, X. Wang, L. Zhou, Y. Tong, R. Ren, K. S. M. Leung, E. H. Y. Lau, J. Y. Wong, X. Xing, N. Xiang, Y. Wu, C. Li, Q. Chen, D. Li, T. Liu, J. Zhao, M. Liu, W. Tu, C. Chen, L. Jin, R. Yang, Q. Wang, S. Zhou, R. Wang, H. Liu, Y. Luo, Y. Liu, G. Shao, H. Li, Z. Tao, Y. Yang, Z. Deng, B. Liu, Z. Ma, Y. Zhang, G. Shi, T. T. Y. Lam, J. T. Wu, G. F. Gao, B. J. Cowling, B. Yang, G. M. Leung, Z. Feng, N. Engl. J. Med. 2020, 382, 1199.

[40] W. A. Knaus, E. A. Draper, D. P. Wagner, J. E. Zimmerman, Crit. Care Med. 1985, 13, 818.

[41] G. Janson, C. Zhang, M. G. Prado, A. Paiardini, Bioinformatics 2017, 33, 444.