



CORRESPONDENCE

Systematic profiling of SARS-CoV-2-specific IgG epitopes at amino acid resolution

Huan Qi¹, Ming-liang Ma¹, He-wei Jiang¹, Jian-ya Ling², Ling-yun Chen³, Hai-nan Zhang¹, Dan-yun Lai¹, Yang Li¹, Zi-wen Guo⁴, Chuan-sheng Hu⁴, Shu-juan Guo¹, Qing-feng Meng¹, Yan Ren⁵, Xiao Yang⁶, Wei Wang⁷, Jie Zhou⁶, Xiao-dong Zhao¹, Hua Li⁴ and Sheng-ce Tao¹

Cellular & Molecular Immunology (2021) 18:1067–1069; <https://doi.org/10.1038/s41423-021-00654-3>

There is a worldwide pandemic of COVID-19 caused by SARS-CoV-2.¹ By January 28, 2021, more than 100 million cases had been diagnosed, and more than 2 million deaths had been reported (<https://coronavirus.jhu.edu/map.html>). The comprehensive and in-depth elucidation of SARS-CoV-2-specific IgG responses will help us to better understand COVID-19 immunity and facilitate the precise development of neutralizing antibodies and vaccines. To date, IgG responses at the protein and peptide levels,^{2–4} but not at the amino acid level, have been at least partially revealed by these studies. Herein, we aimed to go one step further to dissect SARS-CoV-2-specific IgG responses systematically at the amino acid level. We adopted AbMap,⁵ a method that was developed recently for high-throughput epitope mapping. A set of 55 convalescent sera and 226 protein/peptide-enriched antibodies from these sera were analyzed. A map of SARS-CoV-2-specific IgG epitopes at amino acid resolution was generated for the first time.

We collected sera from 55 convalescent patients and 25 controls (Table S1). The patients were hospitalized at Foshan Fourth Hospital in China from 2020-1-25 to 2020-3-8 for various durations. The 55 convalescent sera were subjected to AbMap analysis directly. We identified 418 motifs (Table S2), 275 of which could be matched to 27 of the 28 known SARS-CoV-2 proteins (Table S3, Fig. S1A). Viewing the data vertically (Fig. S1A), the results clearly showed that there were significant differences among the patients. The number of patients was inversely proportional to the frequency (Fig. S1B). Viewing the data horizontally, there were large variations in sum_epitopes and the ratio of sum_epitopes to length of proteins (Fig. S1C).

To reveal more SARS-CoV-2 epitopes of high confidence, we enriched the protein-specific antibodies from each sample in a consequential manner, *i.e.*, RBD, S1, S2, N (Fig. S2) and Orf9b. A variety of epitopes were identified at high confidence and could be matched to the sequences of the corresponding proteins (Tables S2 and S3).

We plotted the epitopes and the frequencies alongside the linear sequence and domains of the S protein (Fig. 1A). Two areas that were rich in significant epitopes were identified. The first area covers almost the entire CTD (C-terminal domain), and the second area covers the S2' protease cleavage site and the fusion peptide (FP). Significant epitopes were also identified at the cytoplasmic C-terminal end of the

S protein. These results are highly consistent with a peptide microarray-based study, in which ~4000 samples were analyzed against a set of 197 peptides that covers the entire S protein.^{6,7} In addition, we took the S protein as an example to demonstrate the necessity of enriching SARS-CoV-2-specific antibodies from sera (Fig. S3). We defined epitopes with a frequency ≥ 3 as significant epitopes and obtained 28 epitopes (Fig. S4A); we mapped them to the 3D structure of the S protein⁸ monomer (Fig. 1B) and trimer (Fig. S4B). It is clear that most of these epitopes are located on the surface of the S protein monomer and trimer (Fig. 1C, D and Table S4). High homology was observed among SARS-CoV-2, SARS-CoV and BtCoV-RaTG13, especially between SARS-CoV-2 and SARS-CoV (Fig. 1E and Table S4).

Biotinylated RBD was also applied to enrich specific antibodies from COVID-19 convalescent sera (Fig. S1). One relatively significant epitope was identified (Fig. S5A). To visualize the conformational location of this epitope, we mapped the critical epitope residues (yellow) to the 3D structure of the RBD (Fig. S5B). It is clear that the critical epitope residues are not on the exact binding interface of RBD and ACE2 but adjacent to it (Fig. S5C), which was similar to the results reported for several well-studied antibodies (Fig. S5D–J). Antibodies targeting this epitope may cause conformational changes or spatial hindrance, thus interfering with the binding of ACE2 and these neutralizing antibodies to the RBD.

To explore the roles of the critical epitope residues, we matched the mutation sites with the epitope map of the S protein (Fig. 1A) and found that six mutation sites are also critical epitope residues (Fig. S6A and Table S5), of which D936Y and P1263L could cause 60–70% and 98–99% infectivity loss, respectively.⁹ These mutation sites were also of high frequency (Fig. S6B). Mutations with high frequency, *i.e.*, D614G,¹⁰ were not identified as critical epitope residues. Furthermore, most of the identified critical residues have a low frequency of mutation. Overall, it is reasonable to argue that these critical residues, especially D936 and P1263, may be useful for vaccine development.

To present the distribution of the epitopes on the N protein, we also matched the epitopes to the N protein sequence. The epitopes were distributed across but not evenly on the N protein. One area rich in significant epitopes was readily identified (Fig. S7A), and 6 significant epitopes were identified when the

¹Shanghai Center for Systems Biomedicine, Key Laboratory of Systems Biomedicine (Ministry of Education), Shanghai Jiao Tong University, Shanghai, China; ²State Key Laboratory of Microbial Technology, Shandong University, Qingdao, Shandong, China; ³BGI Education Center, University of Chinese Academy of Sciences, Shenzhen, Guangdong, China; ⁴State Key Laboratory for Oncogenes and Bio-ID Center, School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai, China; ⁵BGI-Shenzhen, Shenzhen, Guangdong, China; ⁶Foshan Fourth People's Hospital, Foshan, Guangdong, China and ⁷Institute of Biophysics, Chinese Academy of Sciences, Beijing, China
Correspondence: Xiao-dong Zhao (xiaodongzhao@sjtu.edu.cn) or Hua Li (kaikaixinxin@sjtu.edu.cn) or Sheng-ce Tao (taosc@sjtu.edu.cn)
These authors contributed equally: Huan Qi, Ming-liang Ma, He-wei Jiang, Jian-ya Ling, Ling-yun Chen

Received: 28 January 2021 Accepted: 5 February 2021
Published online: 4 March 2021

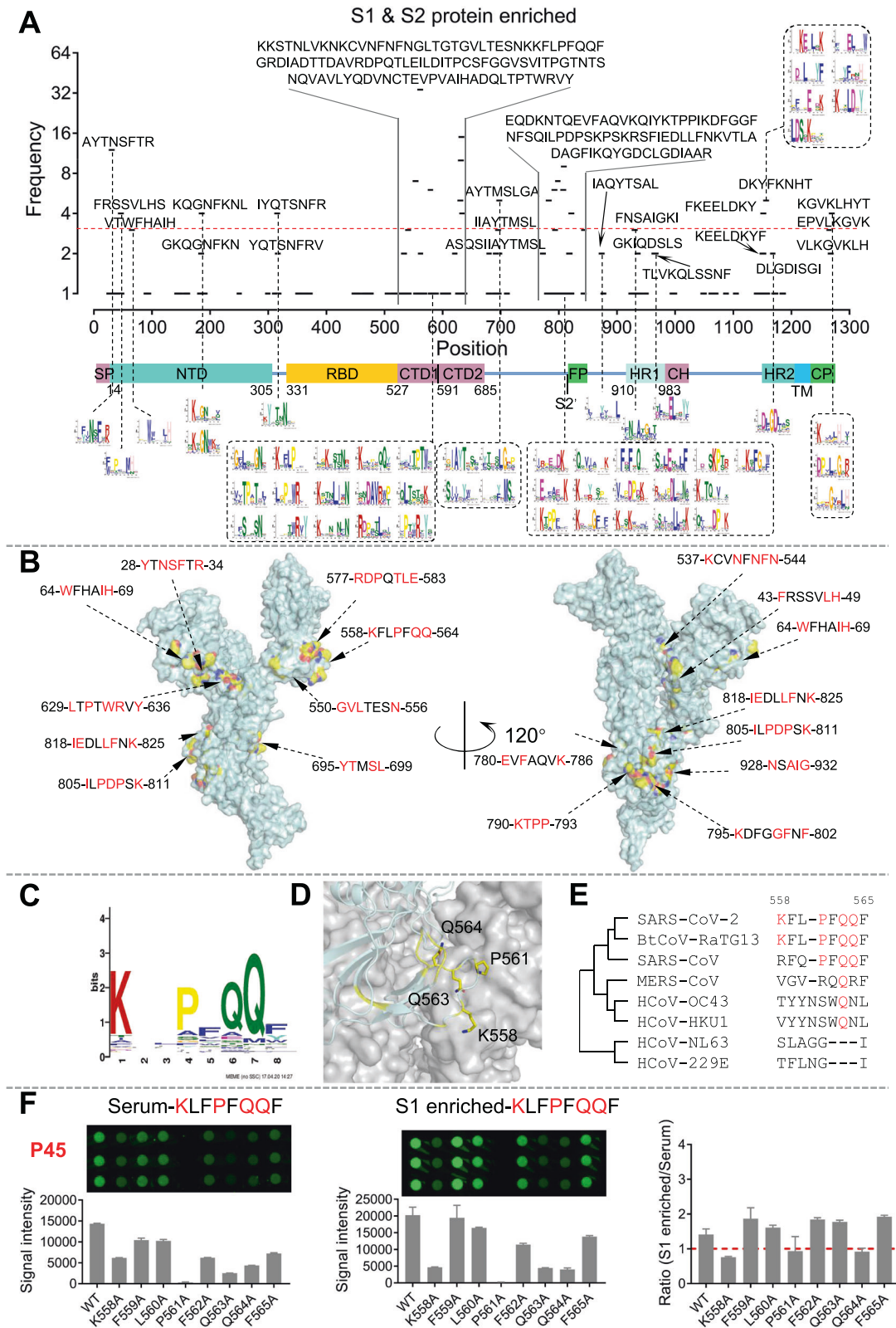


Fig. 1 Significant epitopes on the S protein identified by AbMap and validated by peptide microarray. **A** The distribution of the sequence-matched epitopes on the S protein. **B** The distribution of the significant epitopes (frequency ≥ 3) on the 3D structure of the S protein monomer. Red amino acids represent key residues of epitopes. **C** A representative significant epitope. **D** The epitope was matched to the S protein, and critical residues were labeled yellow. **E** Homology analysis of the epitope among coronaviruses. **F** Validation of epitope identification procedures by a peptide microarray: comparison of sera and protein-enriched antibodies. Peptide 1, which corresponds to Epitope-S1-8 (KLFPFQQF), was selected as an example. The ratio of signal intensity (S1 enriched)/signal intensity (serum) for each residue was plotted (right)

cutoff was set as frequency ≥ 3 (Fig. S7B). High homology for the epitopes and the critical epitope residues was observed among SARS-CoV-2, SARS-CoV and BtCoV-RaTG13 for all 6 significant epitopes (Fig. S7C).

To validate the key residues of the significant epitopes, we selected four epitopes, i.e., Epitope-S1-8 and Epitope-S1-12, which are of the highest frequency, Epitope-RBD-1 from the RBD, and an epitope of low frequency, Epitope-S2-15, as a control. We then mutated the amino acids of these epitopes to alanine one by one, synthesized the mutated peptides, conjugated them to BSA, and fabricated a peptide microarray (Fig. S8). To validate the epitopes on the peptide microarray, we selected 5 sera, from which at least one of the significant epitopes among Epitope-S1-8, Epitope-S1-12 and Epitope-RBD-1 was identified by AbMap (Fig. S9A). For all three epitopes, when the corresponding samples were tested, significant binding signal loss was observed when any of the critical residues was mutated to alanine (Figs. 1F and S9B). In addition, the results in Figs. 1F and S10 further confirmed the necessity for the enrichment of antibodies from sera by specific proteins for epitope mapping.

The aim of this study was to promote our understanding of the SARS-CoV-2-specific IgG response in a more precise way. Currently, no traditional platforms can provide epitope information at the amino acid level. Taking advantage of AbMap, we constructed the first map of SARS-CoV-2-specific IgG binding epitopes at amino acid resolution. We identified several critical epitope residues on the S protein that are highly correlated with the infectivity of SARS-CoV-2 and the binding of ACE2/neutralizing antibodies to the S protein.

There are a variety of possible applications for the significant epitopes/critical residues identified in this study. For example, researchers can select one or a few epitopes based on neutralization activity to generate site-specific antibodies or vaccines on demand for S protein as well as other proteins. Since the COVID-19 pandemic is still unfolding and the sequenced genomes of SARS-CoV-2 are accumulating, undoubtedly, more functionally important amino acid sites will be identified, and we can match those sites to the identified critical epitope residues, thus linking mutation or SARS-CoV-2 evolution with immune responses.

Taken together, these results facilitate the in-depth understanding of SARS-CoV-2-specific IgG responses and provide a basis for the precise development of diagnostic reagents, therapeutic antibodies and even vaccines.

ACKNOWLEDGEMENTS

This work was partially supported by National Key Research and Development Program of China Grant (No. 2016YFA0500600), National Natural Science Foundation of China (Nos. 31970130, 31670831, 31370813, and 31501054), and Open Foundation of Key Laboratory of Systems Biomedicine (No. KLSB2020QN-05).

AUTHOR CONTRIBUTIONS

S.-c.T., X.-d.Z., and H.L. conceived the idea and designed the study; H.Q. performed epitope mapping of sera and the protein-enriched antibodies; M.-l.M. performed the experiments related to the peptide microarray; H.-w.J., L.-y.C., and H.-n.Z. isolated SARS-CoV-2-specific antibodies from sera; Z.-w.G., C.-s.H., X.-d.Z., and H.L. performed NGS and the data analysis; and J.-y.L., L.-y.C., X.Y., W.W., and J.Z. provided key reagents. All the authors analyzed the data, and S.-c.T., H.Q., and M.-l.M. wrote the manuscript. All the authors have read and approved the final manuscript.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41423-021-00654-3>.

Competing interests: The authors declare no conflicts of interest.

REFERENCES

1. Wu, F. et al. A new coronavirus associated with human respiratory disease in China. *Nature* **579**, 265–269 (2020).
2. Jiang, H. W. et al. SARS-CoV-2 proteome microarray for global profiling of COVID-19 specific IgG and IgM responses. *Nat. Commun.* **11**, 3581 (2020).
3. Wang, H. et al. SARS-CoV-2 proteome microarray for mapping COVID-19 antibody interactions at amino acid resolution. *ACS Cent. Sci.* **6**, 2238–2249 (2020).
4. Zamecnik, C. R. et al. ReScan, a multiplex diagnostic pipeline, pans human sera for SARS-CoV-2 antigens. *Cell Rep. Med.* **1**, 100123 (2020).
5. Qi, H. et al. Antibody binding epitope Mapping (AbMap) of hundred antibodies in a single run. *Mol. Cell Proteomics*. <https://doi.org/10.1074/mcp.RA120.002314> (2020).
6. Li, Y. et al. Linear epitopes of SARS-CoV-2 spike protein elicit neutralizing antibodies in COVID-19 patients. *Cell Mol. Immunol.* **17**, 1095–1097 (2020).
7. Li, Y. et al. Systematic evaluation of IgG responses to SARS-CoV-2 spike protein-derived peptides for monitoring COVID-19 patients. *Cell Mol. Immunol.* <https://doi.org/10.1038/s41423-020-00612-5> (2021).
8. Herrera, N. G. et al. Characterization of the SARS-CoV-2 S protein: biophysical, biochemical, structural, and antigenic analysis. *ACS Omega* **6**, 85–102 (2021).
9. Li, Q. et al. The impact of mutations in SARS-CoV-2 spike on viral infectivity and antigenicity. *Cell* **182**, 1284–1294 (2020).
10. Zhang, L. et al. SARS-CoV-2 spike-protein D614G mutation increases virion spike density and infectivity. *Nat. Commun.* **11**, 6013 (2020).