

1 **Antibody repertoire induced by SARS-CoV-2 spike protein immunogens**

2

3 **Supriya Ravichandran, Elizabeth M. Coyle, Laura Klenow, Juanjie Tang, Gabrielle Grubbs,**

4 **Shufeng Liu, Tony Wang, Hana Golding, and Surender Khurana***

5

6 Division of Viral Products, Center for Biologics Evaluation and Research (CBER), FDA, Silver

7 Spring, Maryland, USA, 20871.

8

9 **One Sentence Summary:** SARS-CoV-2 Spike induced immune response

10

11 **Keywords:** SARS-CoV-2, Vaccine, Spike, Neutralization, Epitope, Virus, Antibody, Antigen.

12

13 ***Corresponding author:**

14 *Surender Khurana, Ph.D.

15 Division of Viral Products, Center for Biologics Evaluation and Research (CBER)

16 Food and Drug Administrationa (FDA)

17 10903 New Hampshire Avenue

18 Silver Spring, MD, USA 20993

19 Phone- 240-402-9632, Fax- (301) 595-1125

20 E. mail- Surender.Khurana@fda.hhs.gov

21 **ABSTRACT**

22 Multiple vaccine candidates against SARS-CoV-2 based on viral spike protein are under
23 development. However, there is limited information on the quality of antibody response generated
24 following vaccination by these vaccine modalities. To better understand antibody response
25 induced by spike protein-based vaccines, we immunized rabbits with various SARS-CoV-2 spike
26 protein antigens: S-ectodomain (S1+S2) (aa 16-1213), which lacks the cytoplasmic and
27 transmembrane domains (CT-TM), the S1 domain (aa 16-685), the receptor-binding domain
28 (RBD) (aa 319-541), and the S2 domain (aa 686-1213 as control). Antibody response was analyzed
29 by ELISA, Surface Plasmon Resonance (SPR) against different Spike proteins in native
30 conformation, and a pseudovirion neutralization assay to measure the quality and function of the
31 antibodies elicited by the different Spike antigens. All three antigens (S1+S2 ectodomain, S1
32 domain, and RBD) generated strong neutralizing antibodies against SARS-CoV-2. Vaccination
33 induced antibody repertoire was analyzed by SARS-CoV-2 spike Genome Fragment Phage
34 Display Libraries (SARS-CoV-2 GFPDL), which identified immunodominant epitopes in the S1,
35 S1-RBD and S2 domains. Furthermore, these analyses demonstrated that surprisingly the RBD
36 immunogen elicited a higher antibody titer with 5-fold higher affinity antibodies to native spike
37 antigens compared with other spike antigens. These findings may help guide rational vaccine
38 design and facilitate development and evaluation of effective therapeutics and vaccines against
39 COVID-19 disease.

40

41 INTRODUCTION

42 The ongoing pandemic of SARS-CoV-2 has resulted in more than 2 million human cases
43 and 125,000 deaths as of 15th April 2020. Therefore, development of effective vaccines for
44 prevention and medical countermeasures for treatment of SARS-CoV-2 infection is a high global
45 priority. The spike glycoprotein has been identified as the key target for protective antibodies
46 against both SARS-CoV-1 and SARS-CoV-2(1-4). Consequently, multiple versions of the SARS-
47 CoV-2 spike proteins are currently under evaluation as vaccine candidates utilizing different
48 modalities and delivery systems(5). However, only limited knowledge exists on antibody
49 repertoire or quality of the immune response generated following vaccination by different spike
50 vaccine antigens. Therefore, it is important to perform comprehensive evaluation of post-
51 vaccination antibody response to elucidate the quality of the immune responses elicited by spike-
52 based vaccine candidates to determine immune markers that may predict clinical benefit which
53 can facilitate evaluation of vaccine candidates.

54 To better understand vaccination-induced antibody response, we immunized rabbits with
55 several SARS-CoV-2 spike proteins: the S-ectodomain (S1+S2) (aa 16-1213) lacking the
56 cytoplasmic and transmembrane domains (delta CT-TM), the S1 domain (aa 16-685), the receptor-
57 binding domain (RBD) (aa 319-541), and the S2 domain (aa 686-1213), as a control. Post-
58 vaccination sera were analyzed by Genome Fragment Phage Display Libraries covering the entire
59 spike gene (SARS-CoV-2 GFPDL) to determine the polyclonal antibody epitope repertoire
60 generated following vaccination as previously applied for other diseases(6-10). In addition, we
61 employed several antibody binding assays (ELISA, Surface Plasmon Resonance (SPR) based real-
62 time kinetics assay) (10-12) and an *in vitro* SARS-CoV-2 pseudovirion neutralization assay to
63 measure the quality and function of the antibodies elicited by the different SARS-CoV-2 spike

64 antigens. This study could inform development and evaluation of SARS-CoV-2 vaccines and
65 therapeutics based on the spike glycoprotein.

66

67 **RESULTS:**

68 **Rabbit immunization with SARS-CoV-2 Spike antigens**

69 Most spike-based vaccines currently under development are designed to contain the
70 receptor-binding domain (RBD; aa 319-541) in some form. Therefore, we evaluated four different
71 commercially available SARS-CoV-2 spike protein and subdomains: the Spike S1+S2 ectodomain
72 (aa 16-1213), the S1 domain (aa 16-685), RBD domain (aa 319-541), and the S2 domain (aa 686-
73 1213) as a control, which is devoid of RBD (Fig. 1A, Suppl. Fig. 1). These spike proteins were
74 either produced in HEK 293 mammalian cells (S1 and RBD) or insect cells (S1+S2 ectodomain
75 and S2 domain). The purified S1+S2 ectodomain, the S1 domain, and the RBD proteins retained
76 the functional activity as demonstrated in SPR assay using human ACE2 protein, the SARS-CoV-
77 2 receptor (Fig. 1B). The S1+S2 ectodomain, S1 domain and RBD (black, blue and red binding
78 curves, respectively) demonstrated high-affinity interaction with human ACE2. The control S2
79 domain protein (green curve), lacking the RBD, did not bind to human ACE2, demonstrating
80 specificity of this receptor-binding assay (Fig. 1B).

81 Female New Zealand white rabbits were immunized twice intra-muscularly at a 14-day
82 interval with 50 µg of the purified proteins mixed with Emulsigen Adjuvant. Sera were collected
83 before (pre-vaccination) and after the first and second vaccination and analyzed for binding
84 antibodies in ELISA and SPR, in a pseudovirion neutralization assay, and by GFPDL analysis.

85

86 **Antibody Response following immunization with different Spike antigens**

87 Serial dilutions of post-second vaccination rabbit sera were evaluated for binding of serum
88 IgG to various spike proteins and domains in ELISA (S1+S2; black, S1; blue, RBD; red, and S2;
89 green) (Fig. 1C). Representative titration curves to spike ectodomain (S1+S2) and to the RBD in
90 IgG-ELISA are shown in Suppl. Fig. 2. End-point titers of the serum IgG were determined as the
91 reciprocal of the highest dilution providing an optical density (OD) twice that of the negative
92 control (Fig. 1C). All four immunogens elicited strong IgG binding to the spike ectodomain
93 (S1+S2). Binding to the individual domains (S1, S2, and RBD) was specific, in that sera generated
94 by S2 vaccination bound to S2, but not to S1 or RBD, and vice-versa (Fig. 1C).

95 SPR allows following antibody binding to captured antigens in real-time kinetics, including
96 total antibody binding in resonance units (Max RU) and affinity kinetics (Suppl. Fig. 3). In ELISA,
97 the antigens directly coated in the wells can be partially denatured increasing the likelihood of
98 presenting epitopes that are not seen on the native form of the protein by the polyclonal serum IgG.
99 On the other hand, in our SPR, the purified recombinant spike proteins were captured to a Ni-NTA
100 sensor chip to maintain the native conformation (as determined by ACE2 binding) to allow
101 comparisons of binding to and dissociation from the four proteins. Importantly, the protein density
102 captured on the chip surface is low (200 RU) and was optimized to measure primarily monovalent
103 interactions, so as to measure the average affinity of antibody binding in the polyclonal serum (8,
104 13). Additionally, while ELISA measured only IgG binding, in SPR, all antibody isotypes
105 contributed to antibody binding to the captured spike antigen. In the current study, all rabbit sera
106 contained anti spike antibodies that were at least 86% IgG (data not shown). Serial dilutions of
107 post-vaccination serum were analyzed for binding kinetics with different spike proteins (Suppl.
108 Fig. 3). The spike ectodomain (S1+S2) generated antibodies that predominantly bound to S1+S2

109 (black bar), followed by the S1 protein (blue bar), and 3-fold lower antibody binding to the RBD
110 and the S2 domain (red and green bars, respectively) (Fig. 1D). The S1 domain antigen induced
111 antibodies that bound with similar titers (Max RU values) to the S1+S2, S1 and RBD proteins
112 (black, blue and red bars, respectively), and did not show reactivity to the S2 domain (green bar).
113 However, the antibody reactivity of rabbit anti-S1 serum to S1+S2 domain was 3-fold lower than
114 the antibodies in the rabbit anti-S1+S2 serum. RBD immunization generated similar high-titer
115 antibody binding to S1+S2, S1 and RBD (black, blue and red bars, respectively), (Fig. 1D). In
116 contrast, the S2 domain induced antibodies that primarily bound to homologous S2 antigen (green
117 bars) and only weakly binding to the S1+S2 ectodomain (black bars), and no binding to either S1
118 or RBD (Fig. 1D).

119 Antibody off-rate constants, which describe the fraction of antigen–antibody complexes
120 that decay per second, were determined directly from the serum sample interaction with SARS-
121 CoV-2 spike ectodomain (S1+S2), S1, S2, and RBD using SPR in the dissociation phase only for
122 sensorgrams with Max RU in the range of 20–100 RU (Suppl. Fig. 3) and calculated using the
123 BioRad ProteOn manager software for the heterogeneous sample model as described before(11).
124 These off rates provide additional important information on the affinity of the antibodies following
125 vaccination with the different spike proteins that are likely to have an impact on the antibody
126 function *in vivo*, as was observed previously in studies with influenza virus, RSV and Ebola virus
127 (13-15). Surprisingly, we observed significant differences in the affinities of antibodies elicited by
128 the four spike antigens (Fig. 1E). Specifically, the RBD induced 5-fold higher affinity antibodies
129 (slower dissociation rates) against S1+S2 (black), S1 (blue) and RBD (red) proteins, compared
130 with the post-vaccination antibodies generated by other three immunogens (Fig. 1E).

131 SARS-CoV-2 neutralization was measured using SARS-CoV-2-FBLuc in a single-cycle
132 PsVN assay in Vero E6 cells. The average percent inhibition by post-first and post-second rabbit
133 vaccination are shown in Fig. 1F. Pre-vaccination rabbit sera (Control Rb) did not neutralize
134 SARS-CoV-2 in PsVN assay. Sera generated by S1+S2-ectodomain, S1 and RBD (1:40 dilution)
135 (but not anti-S2) showed 50-60% virus neutralization after a single vaccination, and 93-98% virus
136 inhibition by the post-second vaccination sera (Fig. 1F).

137

138 **Epitope repertoires recognized by antibodies generated against SARS-CoV-2 spike antigens**

139 The constructed SARS-CoV-2 GFPDL contains sequences ranging from 50-1500 bp long
140 from the spike gene (GenBank #MN908947) with $>10^{7.2}$ unique phage clones. The SARS-CoV-2-
141 GFPDL displayed linear and conformational epitopes with random distribution of size and
142 sequence of inserts that spanned the entire spike gene. SARS-CoV-2 GFPDL panning with
143 individual post-second vaccination rabbit sera were conducted as described in Methods. The
144 numbers of IgG-bound SARS-CoV-2 GFPDL phage clones with different serum sample ranged
145 between 2.6×10^4 to 9.8×10^5 /mL (Fig. 2A). Graphical distribution of representative clones with
146 a frequency of ≥ 2 , obtained after affinity selection, and their alignment to the spike protein of
147 SARS-CoV-2 are shown for the four vaccine groups (Fig. 2 B-E). The spike (S1+S2) ectodomain
148 induced diverse antibody response that included strong binding to epitopes in the C-terminal region
149 of the soluble protein spanning the HR2 region (i.e., multiple phage clones with similar inserts).
150 This region may not be highly exposed on the virions or infected cells but is clearly immunogenic
151 in the soluble recombinant spike ectodomain. In addition, the rabbit anti-S1+S2 antibodies bound
152 diverse epitopes spanning the RBD and to a lesser degree to the N-terminal domain (NTD) and the
153 C-terminal region of S1, and the N-terminus of S2, including the fusion peptide (Fig. 2B and Suppl.

154 Table 1). The S1 domain elicited very strong response against the C-terminal region of S1 protein
155 and a diverse antibody repertoire recognizing the NTD and RBD/RBM regions (Fig. 2C and Suppl.
156 Table 1). The recombinant RBD induced high-titer antibodies that were highly focused to the
157 RBD/RBM (Fig. 2E, and Suppl. Table 1). In contrast, the recombinant S2 immunogen after two
158 immunizations in rabbits elicited antibodies primarily targeting the C-terminus of the S2 protein
159 (CD-HR2).

160 All the immunodominant antigenic sites identified by the SARS-CoV-2 GFPDL panning
161 of all 4 immune sera on the spike sequence are shown in Suppl. Fig. 4. Alignment of the sequence
162 with other coronaviruses shows that some of the antigenic sites are >70% conserved among several
163 coronavirus strains isolated from humans and bats, especially those located in the S2 domain
164 (Suppl. Table 1). Structural depiction of these antigenic sites on the SARS-CoV-2 spike (Suppl.
165 Fig. 5; in blue on PDB#6VSB), demonstrated that most of these antigenic sites identified in the
166 current study are surface exposed on the native prefusion spike(2).

167

168 DISCUSSION

169 In this study, we performed an in-depth evaluation of antibody response generated by
170 various SARS-CoV-2 spike antigens that are similar to the vaccine antigens being used in clinical
171 development(5, 16, 17). Bioinformatics approach previously identified 279 potential B-cell
172 epitopes and 48 potential T cell epitopes in the Spike glycoproteins of SARS-CoV viruses, based
173 on human antibody responses to the SARS-CoV-1 infection and the corresponding epitopes in
174 SARS-CoV-2 spike (Grifoni et al. Table 4) (18). We compared the predominant antigenic sites
175 identified by antibodies in our study generated by different spike antigens with the B cell epitopes
176 predicted by Grifoni *et al.*(18). Four of the predicted B epitopes overlapped with the sequences we

177 identified in our GFPDL analysis: aa 287-317 in NTD-RBD overlaps with our antigenic site aa
178 298-363 which is 77% homologous between SARS-CoV-1 and SARS-CoV-2; aa 524-598 and aa
179 601-640, in the C-terminus of S1 overlap with our antigenic site containing aa 548-632 (78.8 %
180 conservation between SARS-CoV-1 and SARS-CoV-2); aa 802-819 in the S2 domain/FP overlaps
181 with our antigenic site aa 768-828 (83% conserved between SARS-CoV-1 and SARS-CoV-2) (Fig.
182 S4 and Suppl. Table 1). The other epitopes identified in our study cover less conserved sequences
183 between the two SARS-CoV viruses that are unique to the SARS-CoV-2 spike and were not
184 identified in the *in-silico* approach by Grifoni et al.

185 Surprisingly, the S2 domain doesn't appear to elicit as many neutralizing antibodies as
186 RBD or S1. Although S2 contains the fusion peptide, it does not appear to be as immunogenic,
187 compared with S1 or RBD, in generating binding antibodies to the intact spike (S1+S2)
188 ectodomain, as observed in both IgG ELISA and SPR. Even though we characterized the purified
189 proteins in various assays, there is a possibility that the structure of the antigens used in the study
190 is different from the corresponding authentic spike protein on the surface of SARS-CoV-2 virion
191 particle.

192 One unexpected finding in this study was the higher affinity of antibodies elicited by the
193 RBD compared with the other spike antigens (S1+S2 ectodomain, S1 and S2 domains). In earlier
194 studies, with vaccines against H7N9 avian influenza we found important correlation between
195 antibody affinity against the hemagglutinin HA1 globular domain and control of virus loads after
196 challenge of ferrets with H7N9 (19). In study of patients recovering from Zika virus (ZIKV)
197 infections, their antibody affinity against ZIKV E-DIII correlated with lower clinical scores(20).
198 In a large randomised clinical trial of IVIG hyper-enriched for influenza virus antibodies (hIVIG),
199 in adults hospitalised with confirmed influenza A or B infections, a statistically significant

200 virological benefit and clinical benefit for patients infected with B strains, directly correlated with
201 stronger antibody affinities of the hIVIG for circulating B strains (14). In a recent longitudinal
202 study of Ebola virus disease survivor, affinity maturation to Ebola virus GP was associated with a
203 rapid decline in viral replication and illness severity in this patient (13). Thus, vaccines that can
204 elicit high affinity antibodies may have a significant advantage for *in-vivo* clinical outcome of
205 SARS-CoV-2 infection and contribute to amelioration of disease in infected individuals.
206 Therefore, in addition to measurements of antibody-binding titers and virus neutralization, this and
207 the previous studies demonstrate the importance of assessments of antibody affinity maturation
208 during SARS-CoV-2 vaccine trials.

209 In summary, our study highlights the need to perform comprehensive analysis of immune
210 response generated following vaccination or SARS-CoV-2 infection to identify biomarkers of
211 protective immunity. In-depth understanding of quantitative and qualitative aspects of immune
212 responses generated by different spike protein vaccine antigens could aid the development and
213 evaluation of effective SARS-CoV-2 therapeutics and vaccines.

214

215 **ACKNOWLEDGEMENTS:**

216 We thank Keith Peden and Marina Zaitseva for their insightful review of the manuscript.

217

218 **Funding:** The antibody characterization work described in this manuscript was supported by FDA
219 intramural grant funds. The funders had no role in study design, data collection and analysis,
220 decision to publish, or preparation of the manuscript.

221 The content of this publication does not necessarily reflect the views or policies of the Department
222 of Health and Human Services, nor does mention of trade names, commercial products, or
223 organizations imply endorsement by the U.S. Government.

224

225 **AUTHOR CONTRIBUTIONS:**

226 **Designed research:** S.K.

227 **Performed research:** S.R., J. T., E.C., L.K., G. G., S.L., T.W., and S.K.

228 **Conducted Animal study:** L.K., G. G., and S.K

229 **SARS-CoV-2 neutralization assays:** S.L., and T.W.

230 **Contributed to Writing:** H.G. and S.K.

231

232 **Declaration of Interests:** The authors declare no competing interests.

233

234 **Materials & Correspondence.** Correspondence and material requests should be addressed to the
235 corresponding author (S.K.).

236

237 **Ethics Statement:** All animal experiments were approved by the U.S. FDA Institutional Animal
238 Care and Use Committee (IACUC) under Protocol #2008-10. The animal care and use protocol
239 meets National Institutes of Health guidelines.

240

241 **FIGURE LEGENDS**

242 **Figure 1: SARS-CoV-2 spike binding and SARS-CoV-2 neutralization by serum antibodies**

243 generated following rabbit immunization with spike antigens. A) Schematic representation of

244 the SARS-CoV-2 spike protein and subdomains. Spike S1+S2 ectodomain (aa 16-1213) lacks the

245 cytoplasmic and transmembrane domains (CT-TM), S1 domain (aa 16-685), RBD domain (aa 319-

246 541), and S2 domain (aa 686-1213), all containing 6x His tag at C-terminus, were produced in

247 either HEK 293 mammalian cells (S1 and RBD) or insect cells (S1+S2 ectodomain and S2

248 domain). (B) Binding of purified proteins to human ACE2 proteins in SPR. Sensorgrams represent

249 binding of purified spike proteins on His-captured chips to 5 µg/mL human ACE2 protein. (C)

250 Anti-spike reactivity of post-immunization rabbit sera. Serial dilutions of post-second vaccination

251 rabbit sera were evaluated for binding to various spike proteins and domains (S1+S2; black, S1;

252 blue, RBD; red, and S2; green) in ELISA. Representative titration curves are shown in Fig. S2.

253 End-point titers of the serum samples were determined as the reciprocal of the highest dilution

254 providing an optical density (OD) twice that of the negative control (no serum was used as negative

255 control). (D) SPR binding of antibodies from rabbits immunized twice with SARS-CoV-2 antigens

256 to spike protein and domains from SARS-CoV-2 (S1+S2; black, S1; blue, RBD; red, and S2;

257 green). Total antibody binding is represented in resonance units in this figure for 10-fold serum

258 dilution. All ELISA and SPR experiments were performed twice and the researchers performing

259 the assay were blinded to sample identity. The variations for duplicate runs of ELISA and SPR

260 were <8% and <5%, respectively. The data shown are average values of two experimental runs.

261 (E) Antibody off-rate constants, which describe the fraction of antigen–antibody complexes that

262 decay per second, were determined directly from the serum/ sample interaction with SARS-CoV-

263 2 spike ectodomain (S1+S2), S1, S2, and RBD using SPR in the dissociation phase only for the

264 sensorgrams with Max RU in the range of 20–100 RU. (F) Virus neutralization titers were
265 measured against SARS-CoV-2-FBLuc in a single-round pseudovirus neutralization assay in
266 triplicates (see Methods). The average percentage inhibition after the first and second vaccination
267 (1:40 serum dilution) for each group are shown. Pre-vaccination rabbit sera also did not neutralize
268 in PsVN assay (Control Rb).

269

270 **Figure 2: Antibody repertoires generated by different SARS-CoV-2 spike antigens.** (A)
271 Number of IgG-bound SARS-CoV-2 GFPDL phage clones using the post-second vaccination
272 rabbit polyclonal sera from the vaccine groups in Fig 1. (B-E) Graphical distribution of
273 representative clones with a frequency of ≥ 2 , obtained after affinity selection, and their alignment
274 to the Spike protein of SARS-CoV-2 are shown for the four vaccine groups: S1+S2 ectodomain
275 (B), S1 (C), S2 domain (D) and S1-Receptor binding domain (RBD) (E). The thickness of each
276 bar represents the frequency of repetitively isolated phage, with the scale shown enclosed in a red
277 box in the respective alignments in each panel. The GFPDL affinity selection data was performed
278 twice. Similar numbers of phage clones and epitope repertoire were observed in both phage display
279 analyses.

280

281 **SUPPLEMENTAL INFORMATION**

282

283 **Figure S1:** Purified SARS-CoV-2 proteins analyzed by SDS-PAGE under reducing and non-
284 reducing conditions.

285

286 **Figure S2:** Anti-spike reactivity of post-vaccination rabbit sera in ELISA.

287

288 **Figure S3:** Steady-state equilibrium analysis of different dilutions of serum antibodies binding to

289 Spike protein by SPR.

290

291 **Figure S4.** Sequence alignment of spike protein from diverse CoV strains.

292

293 **Figure S5.** Structural representation of antigenic sites identified in SARS-CoV-2 using GFPDL

294 analyses.

295

296 **Supplementary Table 1.** Sequence similarity (%) of SARS-CoV-2 spike antigenic sites with other

297 CoV strains.

298

299 **METHODS:**

300

301 **Recombinant CoV Proteins**

302 Recombinant SARS-CoV-2 proteins were purchased from Sino Biologicals (S1+S2
303 ectodomain; 40589-V08B1, S1; 40591-V08H, RBD; 40592-V08H or S2; 40590-V08B).

304 Recombinant purified proteins used in the study were either produced in HEK 293 mammalian
305 cells (S1 and RBD) or insect cells (S1+S2 ectodomain and S2 domain).

306

307 **Rabbit immunization Studies**

308 Female New Zealand white rabbits (Charles River labs) were immunized twice intra-
309 muscularly at 14-days interval with 50 µg of purified proteins mixed with Emulsigen Adjuvant.
310 Sera were collected before (pre-vaccination) and after 1st and 2nd vaccination and analyzed for
311 binding antibodies in ELISA, SPR, neutralization assay and GFPDL analysis.

312

313 **ELISA**

314 96 well Immulon plates were coated with 100 ng/100 µL of recombinant spike protein and
315 protein domains in PBS overnight at 4°C. Starting at a 1:100 dilution, serum samples were serially
316 diluted 1:5 and applied to the protein-coated plate in 10 µL for 1 hr at ambient temperature. Serum
317 samples were assayed in duplicate. Naïve serum samples were assayed along with the experimental
318 samples. After three washes with PBS/0.05% Tween 20, bound antibodies were detected with a
319 donkey anti-rabbit IgG Fc-specific HRP-conjugated antibody (Jackson Immuno Research) After

320 1hr, plates were washed as before and OPD was added for 10min. Absorbance was measured at
321 492 nm. End titer was determined as 2-fold above the average of the absorbance values of the
322 naïve serum samples. The end titer is reported as the last serum dilution that was above this cutoff.

323

324 **Antibody binding kinetics of post-vaccination sera to recombinant SARS-CoV-2 proteins by**
325 **Surface Plasmon Resonance (SPR)**

326 Steady-state equilibrium binding of post-vaccination rabbit polyclonal serum was
327 monitored at 25°C using a ProteOn surface plasmon resonance (BioRad). The purified
328 recombinant Spike proteins were captured to a Ni-NTA sensor chip with 200 resonance units (RU)
329 in the test flow channels. The native functional activity of the Spike proteins was determined by
330 binding to the 5 µg/mL human ACE2 protein.

331 For serum analysis, the protein density on the chip was optimized such as to measure
332 monovalent interactions independent of the antibody isotype. Serially diluted (10-, 20-, 40-, 80-,
333 and 160-fold of freshly prepared sera were injected at a flow rate of 50 µl/min (120 sec contact
334 duration) for association, and disassociation was performed over a 600-second interval. Responses
335 from the protein surface were corrected for the response from a mock surface and for responses
336 from a buffer-only injection. SPR was performed with serially diluted serum of each animal in this
337 study. Antibody isotype analysis for the SARS-CoV-2 spike protein bound antibodies in the
338 polyclonal serum was performed using SPR. Total antibody binding was calculated with BioRad
339 ProteOn manager software (version 3.1). All SPR experiments were performed twice and the
340 researchers performing the assay were blinded to sample identity. In these optimized SPR
341 conditions, the variation for each sample in duplicate SPR runs was <5%. The maximum resonance

342 units (Max RU) data shown in the figures was the RU signal for the 10-fold diluted serum sample.
343 Antibody off-rate constants, which describe the fraction of antigen–antibody complexes that decay
344 per second, are determined directly from the serum/ sample interaction with SARS CoV-2 spike
345 ectodomain (S1+S2), S1, S2, and RBD using SPR in the dissociation phase only for the
346 sensorgrams with Max RU in the range of 20–100 RU and calculated using the BioRad ProteOn
347 manager software for the heterogeneous sample model as described before(11). Off-rate constants
348 were determined from two independent SPR runs.

349

350 **SARS-CoV-2 pseudovirus production and neutralization assay**

351 Human codon-optimized cDNA encoding SARS-CoV-2 S glycoprotein (NC_045512) was
352 synthesized by GenScript and cloned into eukaryotic cell expression vector pcDNA 3.1 between
353 the BamHI and XhoI sites. Pseudovirions were produced by co-transfection Lenti-X 293T cells
354 with pMLV-gag-pol, pFBLuc, and pcDNA 3.1 SARS-CoV-2 S using Lipofectamine 3000. The
355 supernatants were harvested at 48h and 72h post transfection and filtered through 0.45-mm
356 membranes.

357 For neutralization assay, 50 µL of SARS-CoV-2 S pseudovirions were pre-incubated with
358 an equal volume of medium containing serum at varying dilutions at room temperature for 1 h,
359 then virus-antibody mixtures were added to Vero E6 cells in a 96-well plate. After a 3 h incubation,
360 the inoculum was replaced with fresh medium. Cells were lysed 48 h later, and luciferase activity
361 was measured using luciferin-containing substrate.

362

363 **Gene Fragment Phage Display Library (GFPDL) construction**

364 SARS-CoV-2 spike gene was chemically synthesized and used for cloning and
365 construction of phage display libraries. A gIII display-based phage vector, fSK-9-3, was used
366 where the desired polypeptide can be displayed on the surface of the phage as a gIII-fusion protein.
367 Purified DNA containing spike gene was digested with *DNase I* to obtain gene fragments of 100-
368 1000 bp size range and used for GFPDL construction as described previously (6-8). The phage
369 libraries constructed from the SARS-CoV-2 spike gene display viral protein segments ranging in
370 size from 30 to 350 amino acids, as fusion protein on the surface of bacteriophage.

371

372 **Affinity selection of SARS-CoV-2 GFPDL phages with polyclonal rabbit serum**

373 Prior to panning of GFPDL with polyclonal serum antibodies, serum components that
374 could non-specifically interact with phage proteins were removed by incubation with UV-killed
375 M13K07 phage-coated petri dishes (8). Equal volumes of each post-vaccination rabbit serum were
376 used for GFPDL panning. GFPDL affinity selection was carried out in-solution with protein A/G
377 (IgG) specific affinity resin as previously described (6, 7, 9) Briefly, the individual rabbit serum
378 was incubated with the GFPDL and the protein A/G resin, the unbound phages were removed by
379 PBST (PBS containing 0.1 % Tween-20) wash followed by washes with PBS. Bound phages were
380 eluted by addition of 0.1 N Gly-HCl pH 2.2 and neutralized by adding 8 µL of 2 M Tris solution
381 per 100 µL eluate. After panning, antibody-bound phage clones were amplified, the inserts were
382 sequenced, and the sequences were aligned to the SARS-CoV-2 spike gene, to define the fine
383 epitope specificity in the post-vaccination rabbit sera. The GFPDL affinity selection data was
384 performed blindly in a blinded fashion. Similar numbers of bound phage clones and epitope
385 repertoire were observed in the two GFPDL panning.

386

387 **Data Availability**

388 The datasets generated during and/or analyzed during the current study are available from the
389 corresponding author on reasonable request.

390

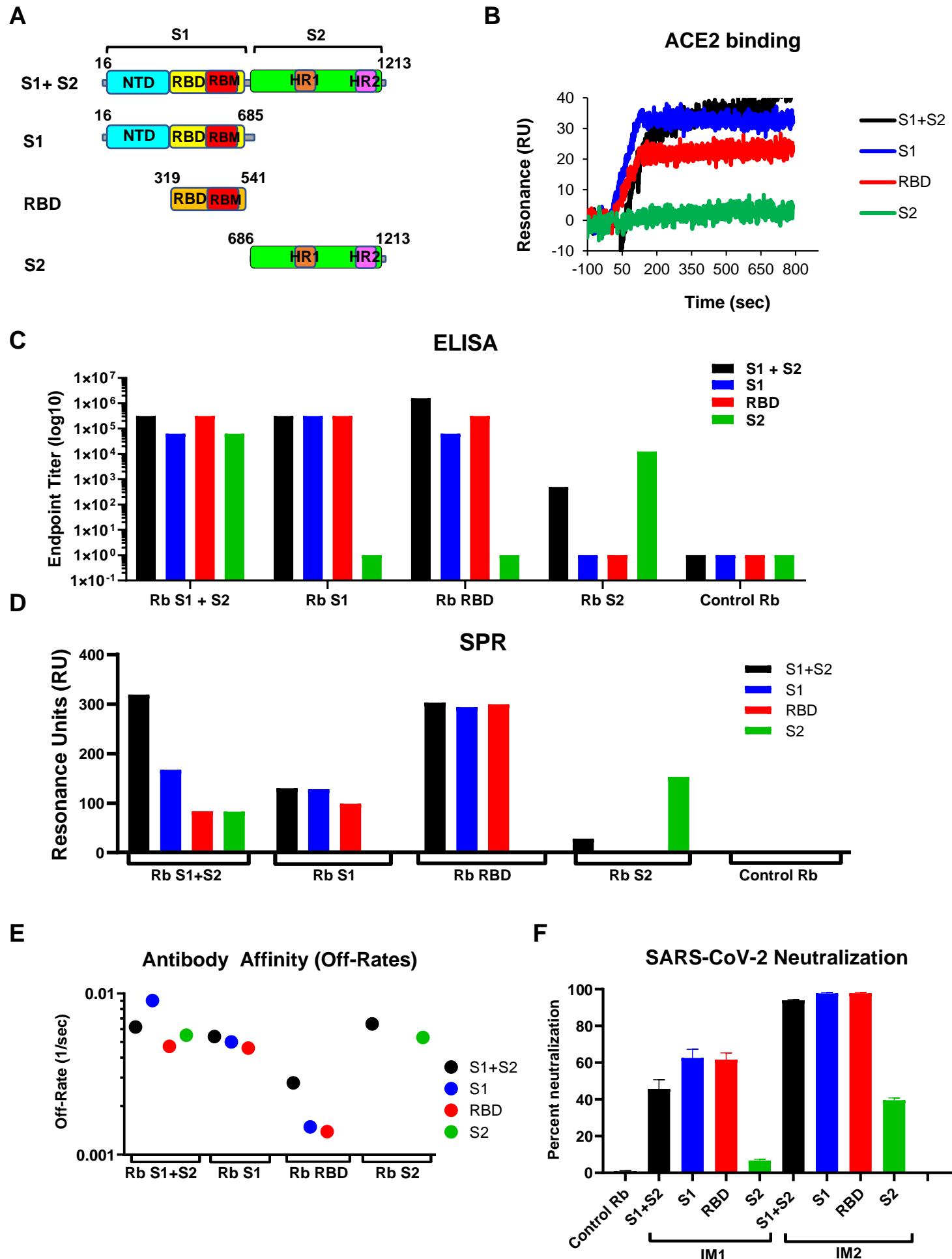
391 REFERENCES

- 392 1. K. Kuba, Y. Imai, S. Rao, H. Gao, F. Guo, B. Guan, Y. Huan, P. Yang, Y. Zhang, W. Deng, L. Bao, B.
393 Zhang, G. Liu, Z. Wang, M. Chappell, Y. Liu, D. Zheng, A. Leibbrandt, T. Wada, A. S. Slutsky, D. Liu,
394 C. Qin, C. Jiang, J. M. Penninger, A crucial role of angiotensin converting enzyme 2 (ACE2) in SARS
395 coronavirus-induced lung injury. *Nature medicine* **11**, 875-879 (2005).
- 396 2. D. Wrapp, N. Wang, K. S. Corbett, J. A. Goldsmith, C. L. Hsieh, O. Abiona, B. S. Graham, J. S.
397 McLellan, Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **367**,
398 1260-1263 (2020).
- 399 3. R. Yan, Y. Zhang, Y. Li, L. Xia, Y. Guo, Q. Zhou, Structural basis for the recognition of SARS-CoV-2
400 by full-length human ACE2. *Science* **367**, 1444-1448 (2020).
- 401 4. A. C. Walls, Y. J. Park, M. A. Tortorici, A. Wall, A. T. McGuire, D. Veesler, Structure, Function, and
402 Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell*, (2020).
- 403 5. F. Amanat, F. Krammer, SARS-CoV-2 Vaccines: Status Report. *Immunity*, (2020).
- 404 6. S. Khurana, A. L. Sugitan, Jr., Y. Rivera, C. P. Simmons, A. Lanzavecchia, F. Sallusto, J.
405 Manischewitz, L. R. King, K. Subbarao, H. Golding, Antigenic fingerprinting of H5N1 avian influenza
406 using convalescent sera and monoclonal antibodies reveals potential vaccine and diagnostic
407 targets. *PLoS Med* **6**, e1000049 (2009).
- 408 7. S. Khurana, W. Clearwae, F. Castellino, J. Manischewitz, L. R. King, A. Honorkiewicz, M. T. Rock,
409 K. M. Edwards, G. Del Giudice, R. Rappuoli, H. Golding, Vaccines with MF59 Adjuvant Expand the
410 Antibody Repertoire to Target Protective Sites of Pandemic Avian H5N1 Influenza Virus. *Science
411 Translational Medicine* **2**, 15ra15-15ra15 (2010).
- 412 8. S. Khurana, S. Fuentes, E. M. Coyle, S. Ravichandran, R. T. Davey, Jr., J. H. Beigel, Human antibody
413 repertoire after VSV-Ebola vaccination identifies novel targets and virus-neutralizing IgM
414 antibodies. *Nature medicine* **22**, 1439-1447 (2016).
- 415 9. S. Fuentes, E. M. Coyle, J. Beeler, H. Golding, S. Khurana, Antigenic Fingerprinting following
416 Primary RSV Infection in Young Children Identifies Novel Antigenic Sites and Reveals Unlinked
417 Evolution of Human Antibody Repertoires to Fusion and Attachment Glycoproteins. *PLoS Pathog*
418 **12**, e1005554 (2016).
- 419 10. S. Khurana, E. M. Coyle, J. Manischewitz, L. R. King, J. Gao, R. N. Germain, P. L. Schwartzberg, J. S.
420 Tsang, H. Golding, C. H. I. C. and the, AS03-adjuvanted H5N1 vaccine promotes antibody diversity
421 and affinity maturation, NAI titers, cross-clade H5N1 neutralization, but not H1N1 cross-subtype
422 neutralization. *NPJ Vaccines* **3**, 40 (2018).
- 423 11. S. Khurana, N. Verma, J. W. Yewdell, A. K. Hilbert, F. Castellino, M. Lattanzi, G. Del Giudice, R.
424 Rappuoli, H. Golding, MF59 adjuvant enhances diversity and affinity of antibody-mediated
425 immune response to pandemic influenza vaccines. *Sci Transl Med* **3**, 85ra48 (2011).
- 426 12. J. L. Halliley, S. Khurana, F. Krammer, T. Fitzgerald, E. M. Coyle, K. Y. Chung, S. F. Baker, H. Yang, L.
427 Martinez-Sobrido, J. J. Treanor, K. Subbarao, H. Golding, D. J. Topham, M. Y. Sangster, High-Affinity
428 H7 Head and Stalk Domain-Specific Antibody Responses to an Inactivated Influenza H7N7 Vaccine
429 After Priming With Live Attenuated Influenza Vaccine. *J Infect Dis* **212**, 1270-1278 (2015).
- 430 13. S. Khurana, S. Ravichandran, M. Hahn, E. M. Coyle, S. W. Stonier, S. E. Zak, J. Kindrachuk, R. T.
431 Davey, Jr., J. M. Dye, D. S. Chertow, Longitudinal Human Antibody Repertoire against Complete
432 Viral Proteome from Ebola Virus Survivor Reveals Protective Sites for Vaccine Design. *Cell Host
433 Microbe* **27**, 262-276 e264 (2020).
- 434 14. R. T. Davey, Jr., E. Fernandez-Cruz, N. Markowitz, S. Pett, A. G. Babiker, D. Wentworth, S. Khurana,
435 N. Engen, F. Gordin, M. K. Jain, V. Kan, M. N. Polizzotto, P. Riska, K. Ruxrungtham, Z. Temesgen, J.
436 Lundgren, J. H. Beigel, H. C. Lane, J. D. Neaton, I. F.-I. S. Group, Anti-influenza hyperimmune

- 437 intravenous immunoglobulin for adults with influenza A or B infection (FLU-IVIG): a double-blind,
438 randomised, placebo-controlled trial. *Lancet Respir Med* **7**, 951-963 (2019).
- 439 15. S. Fuentes, M. Hahn, K. Chilcote, R. F. Chemaly, D. P. Shah, X. Ye, V. Avadhanula, P. A. Piedra, H.
440 Golding, S. Khurana, Antigenic Fingerprinting of Respiratory Syncytial Virus (RSV)-A-Infected
441 Hematopoietic Cell Transplant Recipients Reveals Importance of Mucosal Anti-RSV G Antibodies
442 in Control of RSV Infection in Humans. *J Infect Dis* **221**, 636-646 (2020).
- 443 16. T. Thanh Le, Z. Andreadakis, A. Kumar, R. Gomez Roman, S. Tollefson, M. Saville, S. Mayhew, The
444 COVID-19 vaccine development landscape. *Nat Rev Drug Discov*, (2020).
- 445 17. W. Tai, L. He, X. Zhang, J. Pu, D. Voronin, S. Jiang, Y. Zhou, L. Du, Characterization of the receptor-
446 binding domain (RBD) of 2019 novel coronavirus: implication for development of RBD protein as
447 a viral attachment inhibitor and vaccine. *Cell Mol Immunol*, (2020).
- 448 18. A. Grifoni, J. Sidney, Y. Zhang, R. H. Scheuermann, B. Peters, A. Sette, A Sequence Homology and
449 Bioinformatic Approach Can Predict Candidate Targets for Immune Responses to SARS-CoV-2. *Cell*
450 *Host Microbe* **27**, 671-680 e672 (2020).
- 451 19. S. Khurana, E. M. Coyle, S. Verma, L. R. King, J. Manischewitz, C. J. Crevar, D. M. Carter, T. M. Ross,
452 H. Golding, H5 N-terminal beta sheet promotes oligomerization of H7-HA1 that induces better
453 antibody affinity maturation and enhanced protection against H7N7 and H7N9 viruses compared
454 to inactivated influenza vaccine. *Vaccine* **32**, 6421-6432 (2014).
- 455 20. S. Ravichandran, M. Hahn, P. F. Belaunzaran-Zamudio, J. Ramos-Castaneda, G. Najera-Cancino, S.
456 Caballero-Sosa, K. R. Navarro-Fuentes, G. Ruiz-Palacios, H. Golding, J. H. Beigel, S. Khurana,
457 Differential human antibody repertoires following Zika infection and the implications for
458 serodiagnostics and disease outcome. *Nat Commun* **10**, 1943 (2019).

459

Fig 1

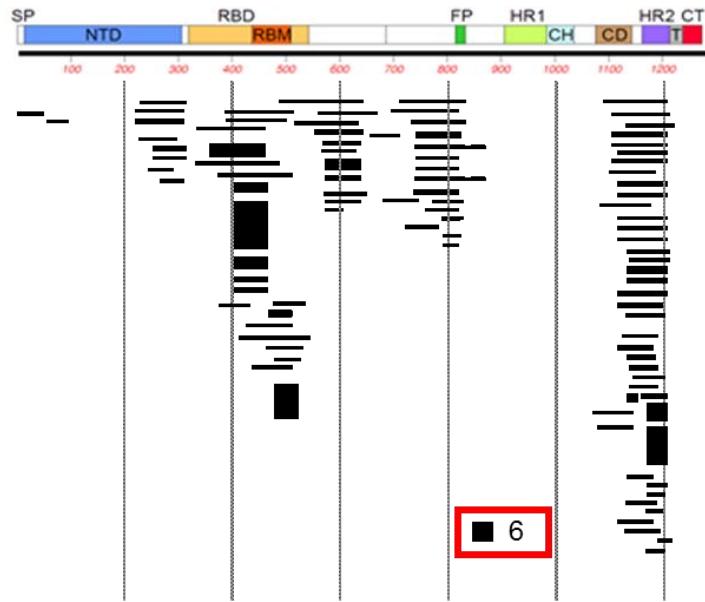


A

	Rb S1+S2	Rb S1	Rb S2	Rb RBD	Control Rb
IgG	5.1E+04	2.2E+05	2.63E+04	9.88E+05	4.20E+01

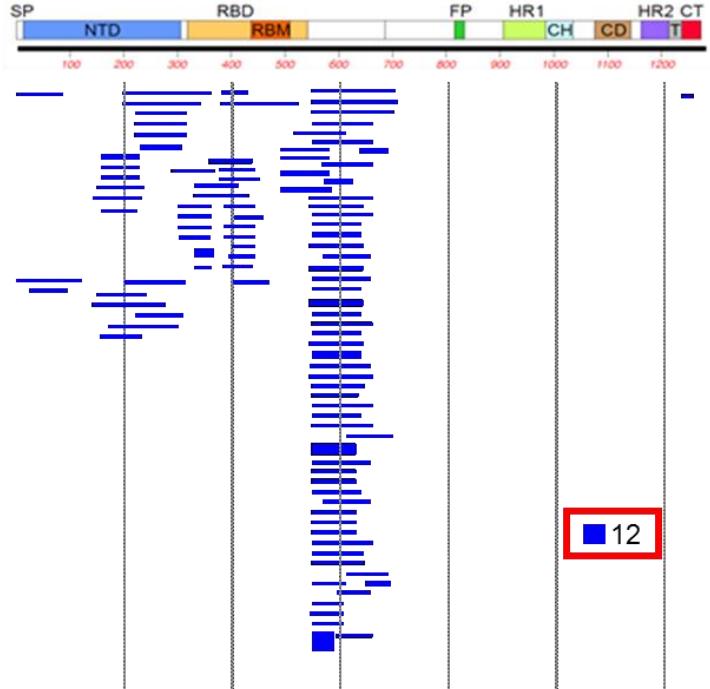
B

S1 + S2 domain (aa 16-1213)



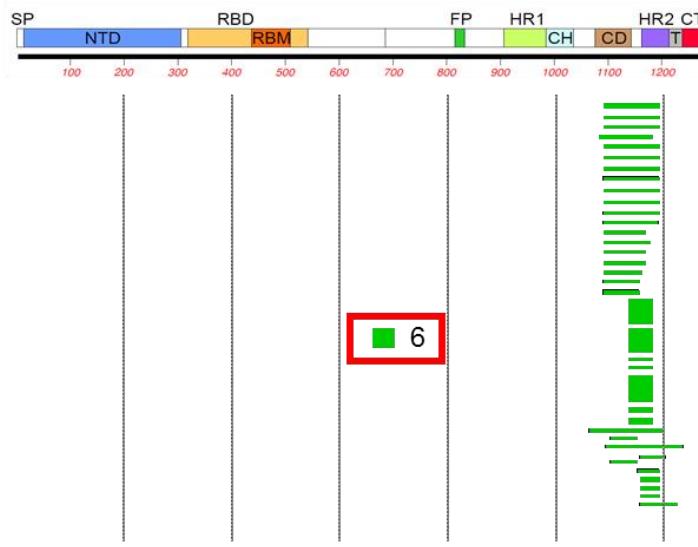
C

S1 domain (aa 16-685)



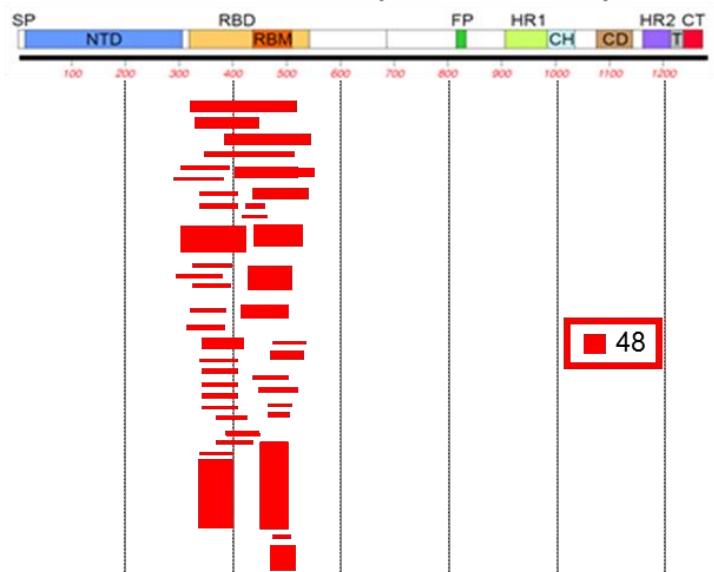
D

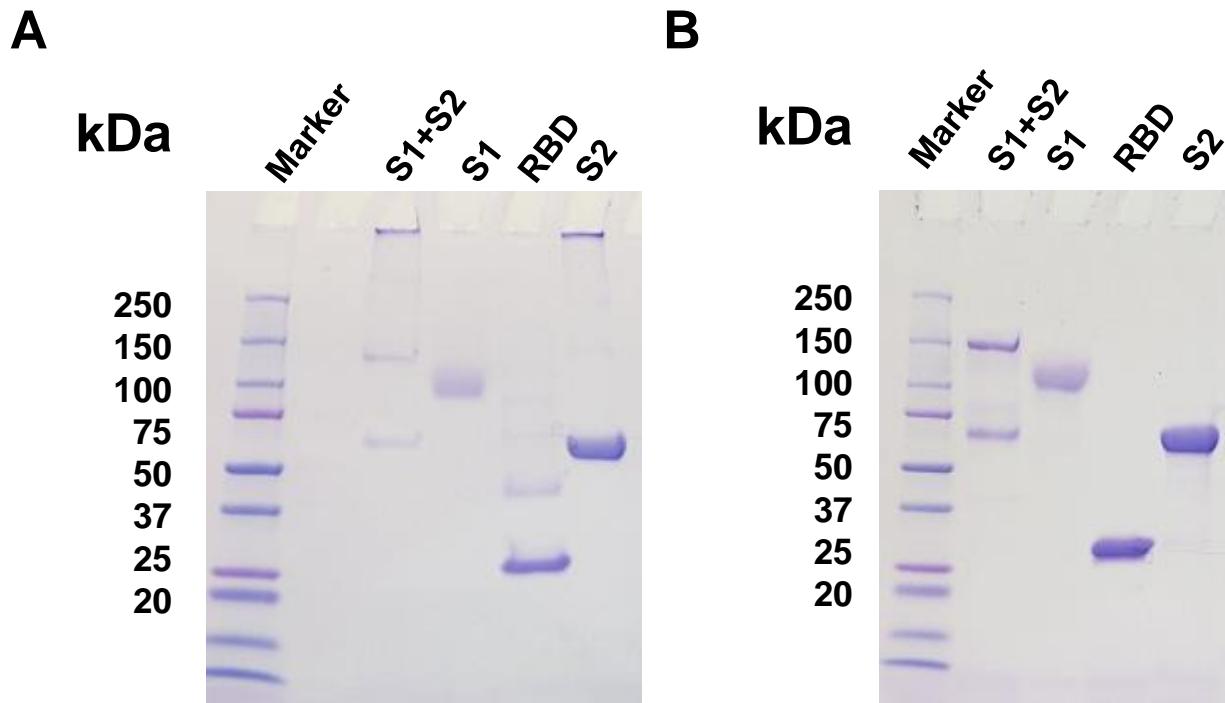
S2 domain (aa 319-541)



E

RBD (aa 686-1213)

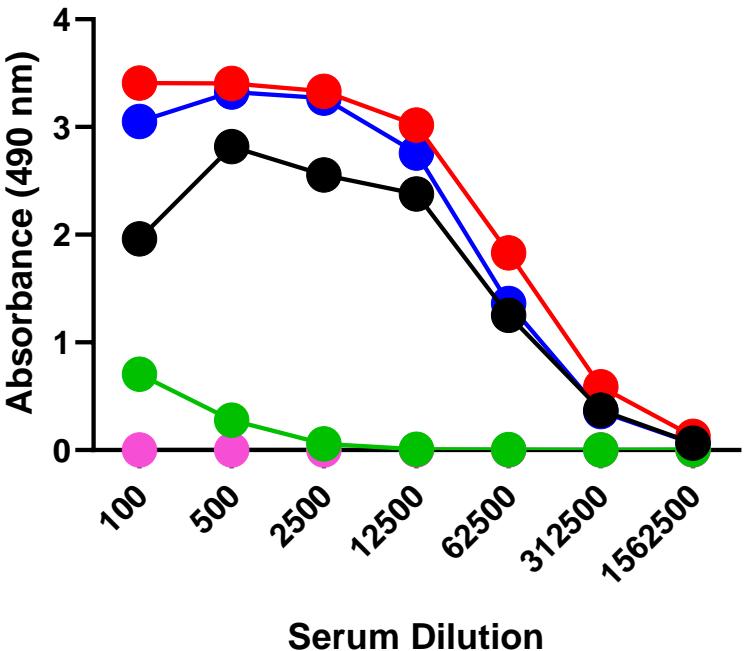




Supplementary Figure 1: Purified SARS-CoV-2 proteins analyzed by SDS-PAGE and under reducing and non-reducing conditions. Related to figure 1. 2 µg of purified proteins was run in SDS-PAGE under non-reducing (A) and reducing (B) conditions. The gels were stained with Coomassie blue.

A

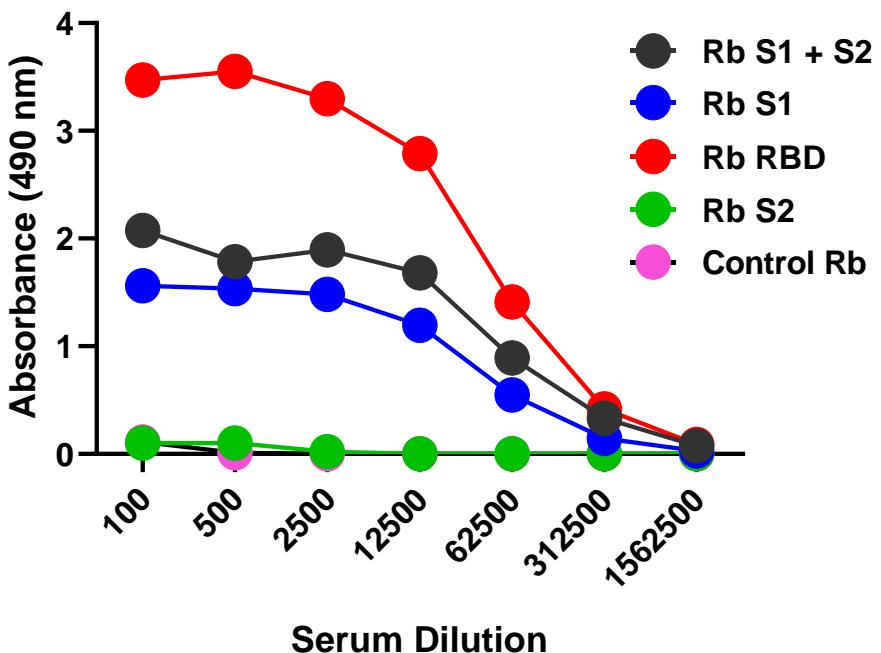
S1 + S2



Serum Dilution

B

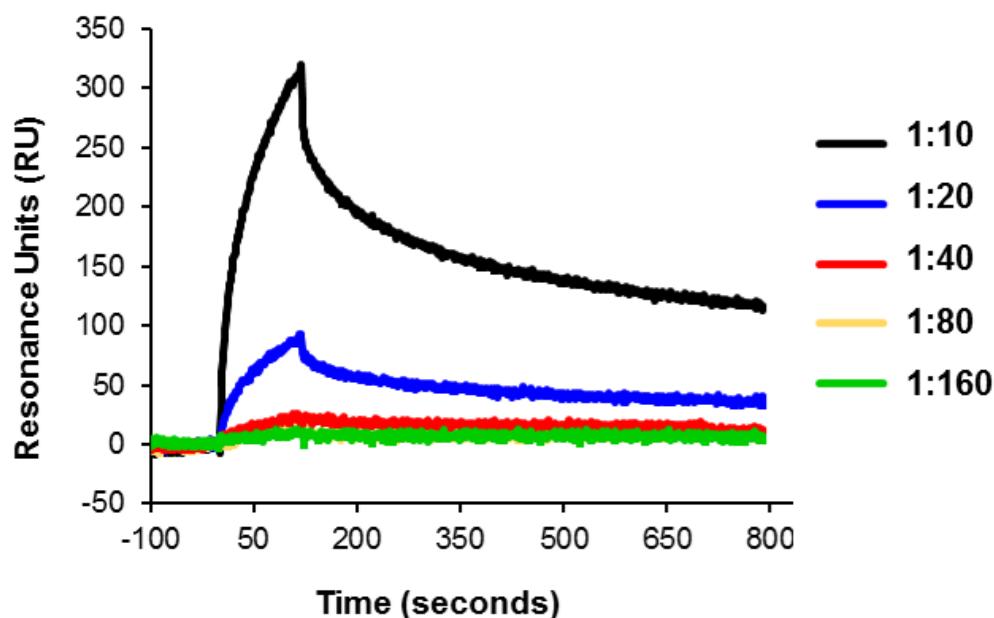
RBD



Serum Dilution

Supplementary Figure 2: Anti-Spike reactivity of post-vaccination rabbit sera in ELISA. Related to figure 1. Post-vaccination rabbit sera following two immunizations with different SARS-CoV-2 spike vaccine antigens (S1+S2; black, S1; red, RBD; blue, S2; green and pre-vaccination control; pink) in ELISA. Average antibody binding to recombinant Spike (S1+S2) ectodomain (A) and S1-RBD (B) is shown in ELISA using HRP-conjugated goat anti-rabbit IgG specific antibody.

Rb S1+ S2



Supplementary Figure 3

Steady-state equilibrium analysis of serum antibodies binding by SPR. Related to figure 1. Serial dilutions of post-2nd vaccination rabbit antiserum against SARS-CoV-2 Spike (S1+S2 ectodomain) were injected simultaneously onto SARS-CoV-1 S1+S2 captured on a Ni-NTA sensor chip and on a surface free of protein (used as a blank). Binding responses from the protein surface were corrected for the response from the mock surface and for responses from a separate, buffer only injection. Unvaccinated Rabbit control sample at 10-fold dilution did not show any binding in SPR.

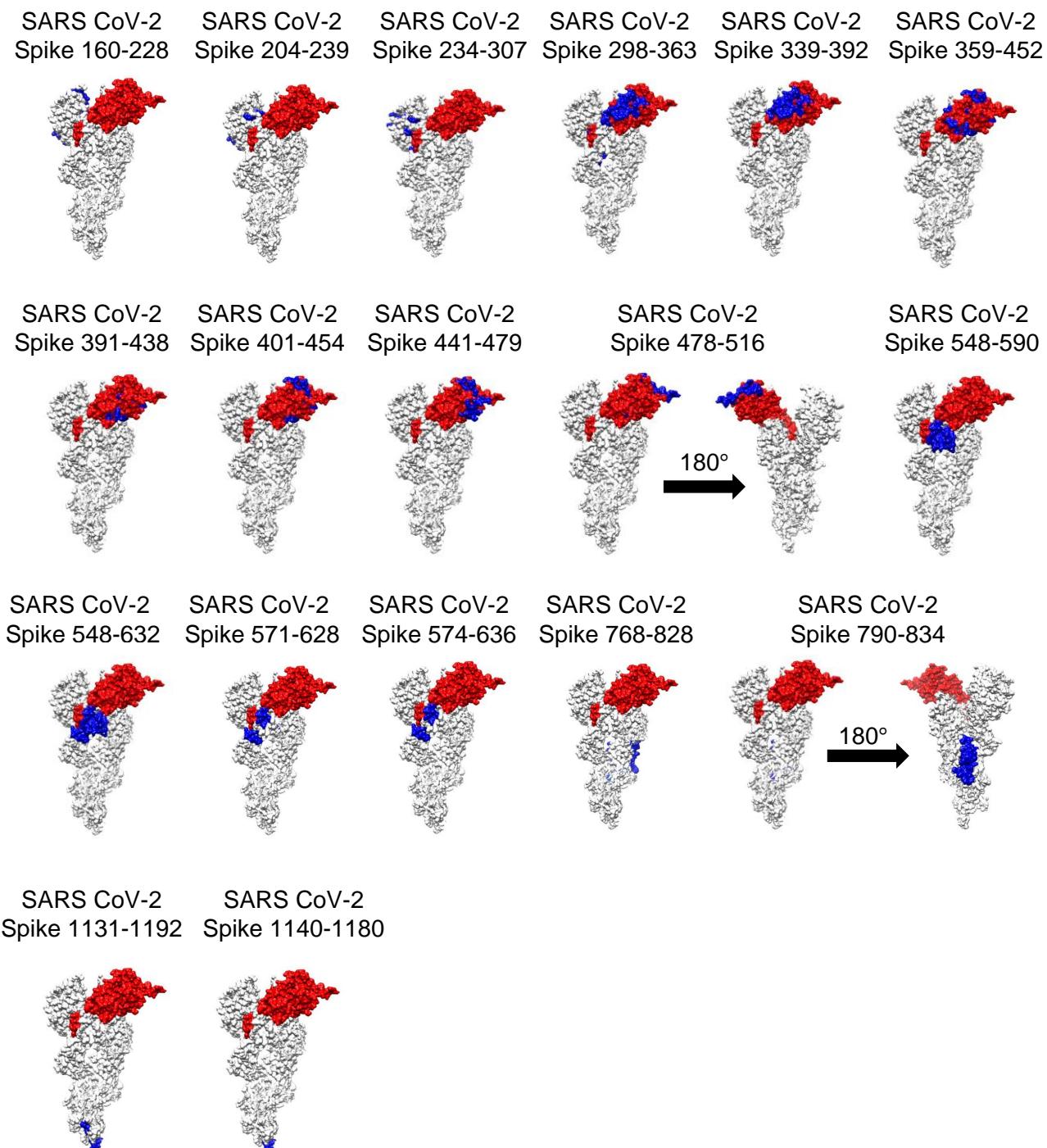
Supplementary Figure 4. Sequence alignment of Spike protein from diverse CoV strains. Related to figures 1 & 2. An alignment of the spike proteins of SARS-CoV-2 (Genbank#MN908947), SARS-CoV-1 BJ01 strain (Genbank#AAP30030.1), MERS CoV KOR/KNIH/2015(Genbank#AKN11075.1), Bat SARS-like CoV ZC45 (Genbank#AVP78031.1), Bat SARS-like CoV ZXC21 (Genban#AVP78042.1), Bat CoV BM48-31/BGR/2008 (Genbank#ADK66841.1), Human CoV 2c EMC/2012 (Genbank#AFS88936.1), Human CoV NL63 (NCBI#YP_003767.1), and Human CoV HKU1 (NCBI#YP_173238.1) was performed using Clustal W multiple alignment application. Various domains of the spike protein are the S1 subunit (AA 1-685), RBD (AA 319-541), FP (816-834) and S2 (686-1273) subunits. The SARS-CoV-2 antigenic regions/sites discovered in this study using the post-vaccination antibodies with different SARS-CoV-2 vaccine antigens are depicted above the SARS-CoV-2 spike protein sequence in alternating black and grey lines with the corresponding AA residues for visualization.

... and is also made available for use under aCC0 license.

SARS CoV-2	--MFVFLVLVPLVS--	-----SQ CVNLTRTQL PPAYTN-----	-SFTRGVYYP DKVFRSSVLH
SARS CoV-1 BJ01	--MFIFLLFLTLTSGSD--	-----LDR CTTFDDVQAF NYTQTN-----	-SSMRGVYYP DEIFRSDTLY
MERS CoV KOR/2015	-MIHSVFLLM FLLTPTESYV DVGPDSVKSA	CIEVDIQQTF FDKTWPRPID VSKADIIYIP	QGRRTYSNITI
Bat SL Cov ZC45	-MLFFFLFLQF ALVN-----	SQ CVNLTRTQL PNPNYTN-----	-SSQRGVYYP DTIYRSDLTV
Bat SL Cov ZXC21	-MLFFFLFLQF ALVN-----	SQ CD-LTGRTPL PNPNYTN-----	-SSQRGVYYP DTIYRSDLTV
Bat CoV BM48-31/2008	-MKFLAFLCL LGFANAQ-----	DGK CGTLSNKSPS KLTQTP-----	-SSRRGFYYP DDIRSSIRV
Human CoV 2c EMC/2015	-MIHSVFLLM FLLTPTESYV DVGPDSVKSA	CIEVDIQQTF FDKTWPRPID VSKADIIYIP	QGRRTYSNITI
Human CoV NL63	MKLFLILLVPLASCFCTCN SANLMSMLQL	GVPDMSSTIV TGLLPTHWFC ANQSTSVD	NGFFYIDVGN
Human CoV HKU1	-MLLIIFILEP TTLAVIG-----	DFNCT NFAINDLNTT VERISEYVVD	VSYGLTYYY LDRVYLNNTI
	80 90 100 110 120 130 140		
SARS CoV-2	STQDLFLPFF SN-----	VT WFHAIHVSGT NGTKRFDNPV LPFNDGVYFA STEKSN-----	
SARS CoV-1 BJ01	LTQDLFLPFY SN-----	VT GFHTINHT FDNPV IPFKDGIYFA ATEKSN-----	
MERS CoV KOR/2015	TYQGLFPYQG DHGDMVYSA GHATGTPQK LFVANYSQDV KQFANGFVVR	I GAAANSTGT VIISPSTSAT	
Bat SL Cov ZC45	LSQGYFLPFY SN-----	VS WYYSLTTN-N AATKRTDNP1 LDFKDGIYFA ATEHSN-----	
Bat SL Cov ZXC21	LSQGYFLPFFY SN-----	VS WYYSLTTN-N AATKRTDNP1 LDFKDGIYFA ATEHSN-----	
Bat CoV BM48-31/2008	LTTGHFLPFN TN-----	LT WYLTLKSN-G KQRIYDNP1 INFQDGVYFG LTEKSN-----	
Human CoV 2c EMC/2015	TYQGLFPYQG DHGDMVYSA GHATGTPQK LFVANYSQDV KQFANGFVVR	I GAAANSTGT VIISPSTSAT	
Human CoV NL63	HRSAFALHTG YYD-----	AN QYYIYVTNEI GLNASVTIKI CKFSRNTTFD FLSNASSSF CIVN-----L	
Human CoV HKU1	LETGYFEKSG AN-----	FR DLSLKGTTL STLWYQKEFL SDFNNNGIFSR VRNTKLYVN-----KT	
	150 160 170 180 190 200 210		
SARS CoV-2	--IIRGWIF GTTLDNSRT-----	QSL LIVNNATNNV IRVCEFQFCN DPFLGVYHK NNKSWMESEF	
SARS CoV-1 BJ01	--VVRGWVF GSTMNNKS-----	QSV IIIINNSTNNV IRACNFELCD NPFFAVSKPM GTQTHMT---	
MERS CoV KOR/2015	IRKIYPAFML GSSVGNFSDG KMGRFFNHTL VLLPDGCCTL	IRAFYCILEP RSGNHCPAGN SYTSFATYHT	
Bat SL Cov ZC45	--IIRGWIF GTTLDNS-----	QSL LIVNNATNNV IRVCEFQFCN DPFLGVYHK NNKSWMESEF	
Bat SL Cov ZXC21	--IVRGWIF GTTLDNS-----	QSL LIVNNATNNV IRVCEFQFCN DPFLGVYHK NNKSWMESEF	
Bat CoV BM48-31/2008	--VFRGWIF GSTLDDNTT-----	QSA VLFNNNGTHIV IDVCNFNC DPMFAVNSGQ PYKTIWY---	
Human CoV 2c EMC/2015	IRKIYPAFML GSSVGNFSDG KMGRFFNHTL VLLPDGCCTL	IRAFYCILEP RSGNHCPAGN SYTSFATYHT	
Human CoV NL63	LTTEQLGAPL GITISGET-----	VRL HLYNVTRTFY VPAAYKLTKL SVKCYFNYS CFSVVNATVT	
Human CoV HKU1	LYSEFSTIVI GSVFINNS-----	YTI VVQPHNGVLE ITACQYTMCE YP-----	204-239
	160-228		
SARS CoV-2	-RVYSSANNCTFEYVSPQFL MDLEGKQGNF KNLREFVFKN IDGYFKIYSK HTPINLVRDL PQGFS-ALE		
SARS CoV-1 BJ01	--IFDNAFNC TFEYISDAFS LDVSEKSGNF KHLREFVFKN KDGFLVYKG YQPIDVVRDL PSGFN-TLK		
MERS CoV KOR/2015	PATDCSDGNY NRNASLNSFK EYFNLRNCTF MYTYNITEDE ILEWFGITQT AQGVHLFSSR YVDLYGGNMF		
Bat SL Cov ZC45	-AVYSSYANC TFEYVSKSFN LNISGNGLF NTLREFVFRN VDGHFKIYSK FTPVNLNRGL PTGLS-VLQ		
Bat SL Cov ZXC21	-AVYSFYANC TFEYVSKSFN LNISGNGLF NTLREFVFRN VDGHFKIYSK FTPVNLNRGL PTGLS-VLQ		
Bat CoV BM48-31/2008	--TSAANC TYHR-AHAFN ISTMNPGKF KHFREHFLKN VDGFLVYHN YEPIDLNSGF PSGFS-VLK		
Human CoV 2c EMC/2015	PATDCSDGNY NRNASLNSFK EYFNLRNCTF MYTYNITEDE ILEWFGITQT AQGVHLFSSR YVDLYGGNMF		
Human CoV NL63	VNVTTHNGRV VNYTVCCDCN GYTDNIFSVQ QDGRIPNGFP FNNWFLLTNG STLVGVSRL YQPLRLTCLW		
Human CoV HKU1	-----HTIC KSKGSSRNES WHFDKSEPLC LFKKNFTYV STDWLWYFHRY QERGTFYAYY ADSGMP-TT	234-307	
	220 230 240 250 260 270 280		
SARS CoV-2	-PFLVDPPIGIN ITRFQTLALL HRSYLTGP-----	DSSSGWTAG AAAYYYVGYLQ PRTFLLKYNE NGTITDAVDC	
SARS CoV-1 BJ01	PIFKLPLGIN ITNFRAILTA FSPAQDT-----	WGTS AAAYFVGYLK PTTFMLKYDE NGTITDAVDC	
MERS CoV KOR/2015	QFATLPVYDT IKYYSIIPHIS IRSIQSD-----	RKA WAAFYVYKLQ PLTFLDDFSV DGYIRRAIDC	
Bat SL Cov ZC45	PFLVEPVISN ITKFRLLTII HRGDPM-----	NNGWTAF SAAFYVYKLQ PRTFMLKYNE NGTITDAVDC	
Bat SL Cov ZXC21	PFLVEPVISN ITKFRLLTII HRGDPM-----	NNGWTAF SAAFYVYKLQ PRTFMLKYNE NGTITDAVDC	
Bat CoV BM48-31/2008	PILKLPFGLN ITYVKAIMTL FSSTQSN-----	FADAD ASAYFVGHLK PLTMLVDFDE NGTIIIDAIDC	
Human CoV 2c EMC/2015	QFATLPVYDT IKYYSIIPHIS IRSIQSD-----	RKA WAAFYVYKLQ PLTFLDDFSV DGYIRRAIDC	
Human CoV NL63	PVPGLKSSTG FVYFNATGSD VNCNGYQHNS VVDVMRYNLN FSANSLDNLK SGVIVFKTLQ YDVLFYCSNS		
Human CoV HKU1	FLFLSLYLGTL LSHYYVLPLT CNAISSN-----	TDNE TLQYWVTPS KRQYLLKFDN RGVITNAVDC	298-363 339-392
	290 300 310 320 330 340 350		
SARS CoV-2	PLVDPPLIGIN ITRFQTLALL HRSYLTGP-----	DSSSGWTAG AAAYYYVGYLQ PRTFLLKYNE NGTITDAVDC	
SARS CoV-1 BJ01	PIFKLPLGIN ITNFRAILTA FSPAQDT-----	WGTS AAAYFVGYLK PTTFMLKYDE NGTITDAVDC	
MERS CoV KOR/2015	QFATLPVYDT IKYYSIIPHIS IRSIQSD-----	RKA WAAFYVYKLQ PLTFLDDFSV DGYIRRAIDC	
Bat SL Cov ZC45	PFLVEPVISN ITKFRLLTII HRGDPM-----	NNGWTAF SAAFYVYKLQ PRTFMLKYNE NGTITDAVDC	
Bat SL Cov ZXC21	PFLVEPVISN ITKFRLLTII HRGDPM-----	NNGWTAF SAAFYVYKLQ PRTFMLKYNE NGTITDAVDC	
Bat CoV BM48-31/2008	PILKLPFGLN ITYVKAIMTL FSSTQSN-----	FADAD ASAYFVGHLK PLTMLVDFDE NGTIIIDAIDC	
Human CoV 2c EMC/2015	QFATLPVYDT IKYYSIIPHIS IRSIQSD-----	RKA WAAFYVYKLQ PLTFLDDFSV DGYIRRAIDC	
Human CoV NL63	PVPGLKSSTG FVYFNATGSD VNCNGYQHNS VVDVMRYNLN FSANSLDNLK SGVIVFKTLQ YDVLFYCSNS		
Human CoV HKU1	FLFLSLYLGTL LSHYYVLPLT CNAISSN-----	TDNE TLQYWVTPS KRQYLLKFDN RGVITNAVDC	398-363 339-392
	360 370 380 390 400 410 420		
SARS CoV-2	ALDPLSETKC TLKSFTVEKG IYQTSNFRVQ PTEISIVRFN ITNLCPFGEV FNATRFASVY AWNRKRISNC		
SARS CoV-1 BJ01	SQNFPLAEILKC SVKSFEIDKG IYQTSNFRVQ PSGDVFVRFN ITNLCPFGEV FNATKFPSVY AWERKKISNC		
MERS CoV KOR/2015	GFNDLSQLHC SYESFDVESG VYSVSSFEAK PSGSVAE QAE G-VECDFSPL LSGT-PPQVY NFKRLVFTNC		
Bat SL Cov ZC45	ALDPLSETKC TLKSFTVEKG IYQTSNFRVQ PTQSVVRFN ITNVCpfHKV FNATRFPSVY AWERTKISDC		
Bat SL Cov ZXC21	ALDPLSETKC TLKSFTVEKG IYQTSNFRVQ PTQSVVRFN ITNVCpfHKV FNATRFPSVY AWERTKISDC		
Bat CoV BM48-31/2008	SQDFPLSELKC TTTSFTVEKG IYQTSNFRVQ PTTEVVRFN ITQLCPFNEV FNITSFPSVY AWERMRTIN		
Human CoV 2c EMC/2015	GFNDLSQLHC SYESFDVESG VYSVSSFEAK PSGSVAE QAE G-VECDFSPL LSGT-PPQVY NFKRLVFTNC		
Human CoV NL63	SSG-VLDTTI PFGPSSQPY CFINSTINT HVSTFVGILP PTVREIVVAR TGQFYINGFK YFDLGIEAV		
Human CoV HKU1	SSSSFFSEIQC KTKSLLPNTG VYDLSGFTVK PVATVHRRIP DLPDCDIDKW LNNFNVPSPL NWERKIFSCN	359-452	
	430 440 450 460 470 480 490		
SARS CoV-2	VADYSVLYN-SASFSTFKCY GVSPTKLNDL CFTNVYADSF VIRGDEVRIQI APGQGTGKIA	YNYKLPDDFT	
SARS CoV-1 BJ01	VADYSVLYN-STFFSTFKCS GVSATKLNDL CFSNVYADSF VVKGDDVRQI APGQGTGVIAD	YNYKLPDDFM	
MERS CoV KOR/2015	NYNLTKLLS-LFSVNDFTCS QISPAAIASN CYSSLILLYF SYPLSMKSDL SVSSAGPISO	FNYKQSFNSNP	
Bat SL Cov ZC45	IADYTVFYN-STSFTFKCY GVSPSKLIDL CFTSVYADTF LIRFSEVRQV APGQGTGVIAD	YNYKLPDDFT	
Bat SL Cov ZXC21	IADYTVFYN-STFSFTFKCY GVSPSKLIDL CFTSVYADTF LIRFSEVRQV APGQGTGVIAD	YNYKLPDDFT	
Bat CoV BM48-31/2008	VADYSVLYN-SASFSTFKCY GVSPTKLNDL CFTSSVYADTF LIRFSEVRQV APGQGTGVIAD	YNYKLPDDFT	
Human CoV 2c EMC/2015	NYNLTKLLS-LFSVNDFTCS QISPAAIASN CYSSLILLYF SYPLSMKSDL SVSSAGPISO	FNYKQSFNSNP	
Human CoV NL63	NFNVTTASAT DFWTVAFATF VDVLVNVSAT NIQNLLYCDS PFEKLQCEHL QFGLQDGFYS ANFLDDNVLP		
Human CoV HKU1	NFNLSTLLR LVHTDSFSCN NFDESKIYGS CFKSIVLDKF AIPNSRRSDL QLGSSGFLQS SNYKIDTTSS		

SARS CoV-2	GCVIAWNSNN	LDSKVGGNYN	YLYRLFR					KSNLKPFERD				
SARS CoV-1 BJ01	GCVLAWNTRN	IDATSTGNYN	YKYRYLR					HGKLRPFERD				
MERS CoV KOR/2015	TCLILATVPH	NLTITKPLK	YSYINKC					SRLLSDDRTE				
Bat SL CoV ZC45	GCVIAWNTAK	QD--VG--N	YFYRSHR					STKLKPFERD				
Bat SL CoV ZXC21	GCVIAWNTAK	QD--TG--H	YFYRSHR					STKLKPFERD				
Bat CoV BM48-31/2008	GCVIAWNTNS	LDSSNE----	FFYRRFR					HGKIKPYGRD				
Human CoV 2c EMC/2015	TCLILATVPH	NLTITKPLK	YSYINKC					SRLLSDDRTE				
Human CoV NL63	ETYVALPIYY	QHTDINFAT	ASFGGSC					-YVCKPHQVN				
Human CoV HKU1	SCQLYYSLPA	INVТИINYNP	SSWNRRYGFN	NFNLLSSHVV	YSRYCFSVNN	TFCPACKPSF	ASSCKSHKPP					
				478-516								
SARS CoV-2	ISTEIQYAGS	TPCNGVEGFN	CY-----	FPLQSYGF	QPTN			GVG YQPYRVVVL				
SARS CoV-1 BJ01	ISNVPFSPDG	KPCTPPALN	CY-----	WPLNDYGF	YTTT			GIG YQPYRVVVL				
MERS CoV KOR/2015	VPQLVNANQY	SPCVSTVPST	VWEDGDYYRK	QLSPLEGGGW	LVAS			GST VAMTEQLQMG				
Bat SL CoV ZC45	LSSDE-----	NGVR		TLSTYDF	NPNV			PLE YQATRVVVL				
Bat SL CoV ZXC21	LSSDE-----	NGVR		TLSTYDF	NPNV			PLE YQATRVVVL				
Bat CoV BM48-31/2008	LSNVLFNPSSG	GTCSAEGLN	CY-----	KPLASYGF	TQSS			GIG FQPYRVVVL				
Human CoV 2c EMC/2015	VPQLVNANQY	SPCVSIVPST	VWEDGDYYRK	QLSPLEGGGW	LVAS			GST VAMTEQLQMG				
Human CoV NL63	ISLNGNTSVC	VRTSHFSIRY	IYN-----	RVKSGSP	GDSS			WHI YLKSGTCPFS				
Human CoV HKU1	SASCPIGTNY	RSCESTVLD	HTDWCR-CSC	LPDPITAYDP	RSCSQKKSLV	GVGEHCAGFG	VDEEKCGVLD					
	640	650	660	670	680	690	700					
SARS CoV-2				FELLHAPATV	CGPKKSTNLV	KN-----		-KCVNFNFNG				
SARS CoV-1 BJ01				FELLNAPATV	CGPKLSTDLI	KN-----		-QCVNFNFNG				
MERS CoV KOR/2015				FGITVQYGTD	TNSVCPKLEF	ANDTKIASQL		GNCVEYSLYG				
Bat SL CoV ZC45				FELLNAPATV	CGPKLSTOLV	KN-----		-QCVNFNFNG				
Bat SL CoV ZXC21				FELLNAPATV	CGPKLSTOLV	KN-----		-KCVNFNFNG				
Bat CoV BM48-31/2008				FELLNAPATV	CGPKQSTELV	KN-----		-KCVNFNFNG				
Human CoV 2c EMC/2015				FGITVQYGTD	TNSVCPKLEF	ANDTKIASQL		GNCVEYSLYG				
Human CoV NL63				FSKLNNFQKF	KTICFSTVEV	PG-----	SCNF	PLEATWHYTS				
Human CoV HKU1				GSYNVSCLCS	TDAFLGWSYD	TCVSNNRCNI	FSNFIINGIN	SGTCSNDLL QPN--TEVFT DVCVDYDLYG				
	548-590						574-636					
SARS CoV-2	LTGTGVLTES	NKKFLP-FQQ	FGRDIADTTD	AVRDPQTLEI	LDITPCSF GG	VSVITPGTNT	SNQVAVLYQD					
SARS CoV-1 BJ01	LTGTGVLTPS	SKRFQP-FQQ	FGRDVSDFTD	SVRDPKTSEI	LDISPCSF GG	VSVITPGTNA	SSEVAVLYQD					
MERS CoV KOR/2015	VSGRGVFQNC	TAVGVR-QQH	FVYDAYQNLV	GYYS-DDGNY	YCLRACVSV	VSVIYD--KE	TKTHATLFGS					
Bat SL CoV ZC45	LKGTVLTD	SKRFQS-FQQ	FGRDASDFID	SVRDPQTLEI	LDITPCSF GG	VSVITPGTNT	SLEVAVLYQD					
Bat SL CoV ZXC21	LKGTVLTD	SKRFQS-FQQ	FGRDASDFID	SVRDPQTLEI	LDITPCSF GG	VSVITPGTNT	SSEVAVLYQD					
Bat CoV BM48-31/2008	LTGTGVLTNS	TKKFQP-FQQ	FGRDVSDFTD	SVRDPKTLEI	LDIAPCSYGG	VSVITPGTNA	SSSVAVLYQD					
Human CoV 2c EMC/2015	VSGRGVFQNC	TAVGVR-QQR	FVYDAYQNLV	GYYS-DDGNY	YCLRACVSV	VSVIYD--KE	TKTHATLFGS					
Human CoV NL63	YTIVGALYVT	WSEGNS-ITG	VPYPVSG-IR	EFSNLVLNNC	TKYNIYDYG	TGIIRSSNQS	LAGGITYVSN					
Human CoV HKU1	ITGQGIFKEV	SAVYYNSWQN	LLYDSNGNII	GFKDFVTNKT	YNIFPCYAGR	VSAAFH--QN	ASSLALLYRN					
	548-632											
SARS CoV-2	VNCTEVPAV	HADQLTPPTWR	VYSTGS--NV	FQTRAGCLIG	AEHV--NNSY	ECDIPIGAGI	CASYQTQTN					
SARS CoV-1 BJ01	VNCTDVTSTAI	HADQLTPAWR	IYSTGN--NV	FQTQAGCLIG	AEHV--DTSY	ECDIPIGAGI	CASYHTVS--					
MERS CoV KOR/2015	VACEHISSTM	SQYSRSTRSM	LKRRDSTYGP	LQTPVGCVLG	LVNSS-LFVE	DCKLPLQSL	CALPDTPSTL					
Bat SL CoV ZC45	VNCTDVPPTI	HADQLTPAWR	IYATGT--NV	FQTQAGCLIG	AEHV--NASY	ECDIPIGAGI	CASYHTAS--					
Bat SL CoV ZXC21	VNCTDVPPTI	HADQLTPAWR	IYAIGT--SV	FQTQAGCLIG	AEHV--NASY	ECDIPIGAGI	CASYHTAS--					
Bat CoV BM48-31/2008	VNCTDVPML	HADQISHDWR	VYAFRNDGNI	FQTQAGCLIG	AAYD--NSSY	ECDIPIGAGI	CAKYTNVSS-					
Human CoV 2c EMC/2015	VACEHISSTM	SQYSRSTRSM	LKRRDSTYGP	LQTPVGCVLG	LVNSS-LFVE	DCKLPLQSL	CALPDTPSTL					
Human CoV NL63	SGNLLGFKNV	STGNIFIVTP	CNQPDQVAVY	QQSII GAMTA	VNESRYGLN	LLQLPNFYVV	SNGNNCTTA					
Human CoV HKU1	LKCSYVLNNI	SLTTQP-----	Y	FDSYLGCVFN	ADNLTDYSVS	SCALRMGSGF	CVDYNSPSS					
	700	750	800	810	820	830	840					
SARS CoV-2	P----RRARS	VASQSIAYT	MSLGAENSA	YSNN--SIA	IPTNFTISVT	TEILPVSM	TSVDCTMYIC					
SARS CoV-1 BJ01	LLRS	TSQKSIVAYT	MSLGAENSA	YSNN--SIA	IPTNFSISIT	TEVMPVSM	TSVDCNMYIC					
MERS CoV KOR/2015	T----PRSVR	SVPGEMLAS	IAFNHPIQVD	QLNSSYFKL	IPTNFSFGV	QEYIQT	TVVDCKQYVC					
Bat SL CoV ZC45	ILRS	TSQKAIVAYT	MSLGAENSA	YANN--SIA	IPTNFSISVT	TEVMPVSM	TSVDCNMYIC					
Bat SL CoV ZXC21	ILRS	TGQKAIVAYT	MSLGAENSA	YANN--SIA	IPTNFSISVT	TEVMPVSM	TSVDCNMYIC					
Bat CoV BM48-31/2008	TLVR	SGGHSILAYT	MSLGDNQDIV	YSNN--TIA	IPMNFSISVT	TEVLPVSM	TSVDCNMYIC					
Human CoV 2c EMC/2015	T----PRSVR	SVPGEMLAS	IAFNHPIQVD	QLNSSYFKL	IPTNFSFGV	QEYIQT	TVVDCKQYVC					
Human CoV NL63	VMTYSNFGIC	ADGSILPVRP	RNSSDNGISA	IITAN--LS	IPSNWTTSVQ	VEYLQITSTP	IVVDCATYVC					
Human CoV HKU1	SSRRKRRSIS	ASYRFVTFEP	FNVSFVNDSI	ESVGGLYEIK	IPTNFTIVGQ	EEFQTNSPK	VTIDCSLFVC					
	768-828											
SARS CoV-2	GDSTECSNL	LQYGSFC	TQLNRALTGIAVE	QDKNTQEVFA	QVKQIYKTPP	IKDFGG-FNF	SQILPDPSK					
SARS CoV-1 BJ01	GDSTECA	NL	NRALSGIAAE	QDRNTR	TEVFA	QVKQMYKTPP	IKDFGG-FNF	SQILPDPLK				
MERS CoV KOR/2015	NGFQKCEQLL	REYQGFC	SKI	NQALHG	ANLR	QDDSVRN	IPGF	GGDFNL	TLLEPV			
Bat SL CoV ZC45	GDSIECSNL	LQYGSFC	TQLNRALTGIAIE	QDKNTQEVFA	QVKQIYKTPP	IKDFGG-FNF	SQILPDPSK					
Bat SL CoV ZXC21	GDSIECSNL	LQYGSFC	TQLNRALTGIAIE	QDKNTQEVFA	QVKQIYKTPP	IKDFGG-FNF	SQILPDPSK					
Bat CoV BM48-31/2008	GDSIECSNL	LQYGSFC	TQLNRALTGIAVE	QDRNTRDVF	QTKAMYKTPS	LKDFFGG-FNF	SQILPDPAK					
Human CoV 2c EMC/2015	NGFQKCEQLL	REYQGFC	SKI	NQALHG	ANLR	QDDSVRN	IPGF	GGDFNL	TLLEPV			
Human CoV NL63	NGNPRCKNL	KOYTSACKTI	EDALRLSAHL	ETNDV	SSMLT	FDSNAFL	SVT	FGD-YNL	SSVLPQRNIR			
Human CoV HKU1	SNYAA	CHDLL	SEYGTFC	DNI	NSILDEV	GL	LDTTQLHVAD	TLMQG	VTLS	DNINF	KSLVG	
	920	930	940	950	960	970	980					

SARS CoV-2 SARS CoV-1 BJ01 MERS CoV KOR/2015 Bat SL CoV ZC45 Bat SL CoV ZXC21 Bat CoV BM48-31/2008 Human CoV 2c EMC/2015 Human CoV NL63 Human CoV HKU1 K-FDEDDSEP VLKGVKLHYT K-FDEDDSEP VLKGVKLHYT DRYEEYDILEP HKVHVH---- K-FDEDDSEP VLKGVKLHYT K-FDEDDSEP VLKGVKLHYT K-FDEDHSEP VLTVVKLHYT DRYEEYDILEP HKVHVH---- DCGSTKLPYY EFEKVHVQ-- DEYGGHHDFV IKTSHDD---
--	--



Supplementary Figure 5. Structural representation of antigenic sites identified in SARS-CoV-2 using GFDPL. Related to figure 2. Antigenic sites identified using GFDPL have been depicted in blue on surface structures of a monomer of PDB#6VSB (Wrapp et al., 2020) with a single receptor-binding domain (RBD) in the up conformation, wherever available using UCSF Chimera software. The RBD region is shaded in red (residues 319-541) on every structure. Those structures (SARS CoV-2 Spike 478-516 and 790-834), whose sites were not visible on the side depicted by flipping the structure by 180°.

Table S1: Sequence conservation of Antigenic regions/sites among different CoV strains

AA	Similarity of SARS CoV 2 antigenic sites to other CoV strains (%)								
	SARS CoV-2	SARS CoV-1	MERS	Bat SL CoV ZC45	Bat SL CoV ZXC21	Bat CoV BM48	Human CoV 2c	Human NL63	Human HKU1
160-228	100	53.6	9.7	62.3	60.8	39.1	9.7	12.6	5.7
204-239	100	55.5	10.5	63.8	63.8	44.4	10.5	21	8.1
234-307	100	52.7	32.4	74.3	71.6	48.6	32.4	6.4	27
298-363	100	77.2	30.3	80.3	80.3	78.7	30.3	7.5	30.3
339-392	100	87	24	75.9	75.9	83.6	24	3.6	27
359-452	100	80.8	20.2	74.4	72.3	76.8	20.2	5.2	20.2
391-438	100	83.3	20.8	83.3	83.3	81.2	20.8	8.3	22.9
401-454	100	74	16.6	72.2	68.5	70.3	16.6	5.5	12.9
441-479	100	46.1	10.2	41	35.8	30.7	10.2	10.2	2.7
478-516	100	61.5	14.2	43.5	43.5	56.4	14.2	12.5	8.7
548-590	100	74.4	16.2	74.4	74.4	79	16.2	4.6	22.7
548-632	100	78.8	17.6	80	80	75.2	17.6	4.7	22
574-636	100	80.9	14.2	87.3	87.3	76.1	14.2	3.1	22.2
768-828	100	83.6	41.5	96.7	96.7	80.3	41.5	24.6	24.2
790-834	100	86.6	46.9	100	100	84.4	46.9	32	31.4
1140-1180	100	100	26.1	97.5	97.5	75.6	26.1	14.5	33.3
1173-1207	100	100	45.7	97.1	97.1	82.8	45.7	20	37.1

* Percent sequence conservation of GFPDL identified antigenic sites in the SARS CoV-2 spike protein (Genbank#MN908947), with CoV Spike proteins of SARS CoV-1 BJ01 strain (Genbank#AAP30030.1), MERS CoV KOR/KNIH/2015(Genbank#AKN11075.1), Bat SARS-like CoV ZC45 (Genbank#AVP78031.1), Bat SARS-like CoV ZXC21 (Genban#AVP78042.1), Bat CoV BM48-31/BGR/2008 (Genbank#ADK66841.1), Human CoV 2c EMC/2012 (Genbank# AFS88936.1), Human CoV NL63 (NCBI#YP_003767.1), and Human CoV HKU1 (NCBI#YP_173238.1) was calculated using Sequence Identity Matrix function in BioEdit.