High-Potency Polypeptide-based Interference for Coronavirus Spike Glycoproteins

Qinghua Wang (✉ qinghuaw@bcm.edu)

Baylor College of Medicine

Jianpeng Ma

Baylor College of Medicine

Adam Acevedo

Baylor College of Medicine

Biological Sciences - Article

Keywords: Coronavirus, COVID-19, inhibitor, SARS-CoV-2, spike glycoproteins

Posted Date: April 5th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-388285/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

1

2 High-Potency Polypeptide-based Interference for Coronavirus Spike Glycoproteins

3

4

5

April 2, 2021

6

7

8

9 **Running Title:** Inhibiting coronavirus spike proteins

10

11

Keywords: Coronavirus, COVID-19, inhibitor, SARS-CoV-2, spike glycoproteins

13

14 **Abstract**

15 The world is experiencing an unprecedented coronavirus disease 2019 (COVID-19) pandemic
16 caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). SARS-CoV-2 spike
17 protein-based vaccines are currently the main preventive agent to fight against the virus. However,
18 several variants with extensive mutations in SARS-CoV-2 spike proteins have emerged. Some of these
19 variants exhibited increased replication, higher transmission and virulence, and were partially resistant
20 to antibody neutralization from natural infection or vaccination. With over 130 million confirmed cases
21 and widespread vaccination around the globe, the emergence of new escape SARS-CoV-2 variants could
22 be accelerated. New therapeutics insensitive to mutations are thus urgently needed. Here we have
23 developed an inhibitor based on SARS-CoV-2 spike protein that potently reduced pseudovirus
24 infectivity by limiting the level of SARS-CoV-2 spike proteins on virion envelope. Most importantly,
25 the inhibitor was equally effective against other coronavirus spike proteins that shared as low as 35%
26 amino-acid sequence identity, underscoring its extreme tolerance to mutations. The small-sized
27 inhibitor would also allow simple delivery by, for instance, nasal spray. We expect the inhibitor
28 reported here to be an invaluable aid to help end COVID-19 pandemic. Furthermore, the use of a partial
29 native sequence or its homologues to interfere with the functions of the native protein represents a novel
30 concept for targeting other viral proteins in combating against important viral pathogens.

31

32 **Main text**

33 In the past 20 years, there have been three major threats from coronavirus, the 2002-2003
34 outbreak in China from severe acute respiratory syndrome coronavirus (SARS-CoV), the 2012-2014
35 outbreak in Middle East from Middle East respiratory syndrome coronavirus (MERS-CoV)^{1,2}, and the
36 ongoing COVID-19 pandemic caused by SARS-CoV-2. Currently, the world is racing to contain

37 SARS-CoV-2 mainly by using SARS-CoV-2 spike protein (SARS2-S)-based vaccines. However,
38 continued genomic evolution has led to SARS-CoV-2 variants with extensive mutations on SARS2-S
39 proteins³⁻⁵ (**Fig.1a**). These mutations endorsed tighter binding with host cell surface receptor, human
40 angiotensin-converting enzyme 2 (hACE2), higher virulence, resistance to antibody neutralization and
41 escape from natural infection or vaccine induced sera⁶⁻¹⁰. Therefore, there is an urgent need for other
42 interventions that are insensitive to genomic mutations.

43 Towards this end, we have developed a polypeptide-based inhibitor, F1, that potently interfered
44 with the formation and translocation of coronavirus spike glycoproteins to host cell surface and to the
45 envelope of generated pseudoviruses. Although derived from the SARS2-S sequence, F1 was also
46 effective on 2002 SARS-CoV spike protein (SARS-S) and 2012 MERS-CoV spike protein (MERS-S).
47 These sequences shared as low as 35% amino-acid sequence identity, stressing the extreme resistance of
48 F1 to mutations in spike sequences of emerging COVID-19 SARS-CoV-2 variants. Different from
49 traditional protein-based therapeutics, the F1 polypeptide-based interference requires that the plasmid or
50 minicircle DNA harboring F1-coding gene is delivered into the host to synthesize F1 polypeptide at the
51 site where the target spike proteins are synthesized. Therefore, the rapid degradation intrinsically
52 associated with protein/polypeptide-based therapeutics is no longer a problem. Moreover, the small-
53 sized F1-minicircle developed here would also allow simple delivery by, for example, nasal spray. We
54 expect the inhibitor reported here to be an invaluable aid to our effort to stop COVID-19 pandemic.

56 **The concept of polypeptide-based protein interference**

57 Viral membrane fusion proteins such as coronavirus spike proteins are oligomeric Class-I
58 transmembrane glycoproteins on the viral envelope¹¹. Coronavirus spike proteins are cleaved to give
59 rise to N-terminal S1 regions and C-terminal S2 regions (**Fig.1a**). The N-terminal S1 regions are the

60 major target for neutralizing antibodies elicited by natural infection or vaccination, and therefore under
61 constant positive selection for escape variants (**Fig.1a**). On the other hand, the C-terminal S2 regions
62 responsible for oligomerization and membrane fusion¹²⁻¹⁶ are more conserved among different
63 coronavirus strains. Upon entry into a host cell, the viral genome will guide the synthesis of new
64 coronavirus spike proteins by ribosomes, which are then folded, assembled and translocated into
65 endoplasmic reticulum (ER) membranes and transit through the ER-to-Golgi intermediate compartment
66 for interaction with newly replicated genomic RNA to produce new virions^{17,18}.

67 We hypothesized that stably foldable fragments of the coronavirus spike proteins, for instance,
68 polypeptides derived from the SARS2-S S2 regions, which maintained the same oligomeric interface as
69 the native spike proteins, would form non-native oligomers with the wild-type spike proteins, thus
70 significantly lowering the level of native spike oligomers on the envelope of newly generated virions,
71 and impairing their infectivity (**Fig.1b**). More importantly, the fragment derived from the conserved S2
72 region would be more resistant to mutations and could be potentially “universal” against different
73 coronaviruses. To test this strategy, we made two polypeptides derived from the SARS2-S sequence, F1
74 and F2, that contained partial SARS2-S S2 sequence (**Fig.1a, Fig.S1, S2**). F1 polypeptide encompassed
75 amino-acid residues 911~1273, while F2 harbored residues 985~1273. Both polypeptides contained a
76 N-terminal signal peptide (SP) as the wild-type SARS2-S for cell surface translocation.

77 **F1 polypeptide potently inhibited expression and cell surface translocation of multiple coronavirus**
78 **spike glycoproteins**

79 We first assessed the impacts of F1 or F2 on the expression and cell surface translocation of
80 SARS2-S protein. Transient transfection of SARS2-S-harboring plasmid yielded a good level of
81 proteins detected in HEK293T whole cell lysate, with most of the expressed proteins were cleaved

82 (Fig.2a). This high-efficiency cleavage of SARS2-S proteins agreed with the novel polybasic cleavage
83 site at the S1/S2 boundary^{12,13,19,20}. Moreover, only the S2 fragments of the cleaved SARS2-S protein
84 were labeled by biotin, affinity-purified by anti-biotin antibody and detected in cell surface fraction
85 (Fig.2a), suggesting that only properly cleaved SARS2-S protein were translocated to cell surface.
86 Most impressively, when F1-harboring plasmid was co-transfected with SARS2-S-carrying plasmid,
87 even at a twofold molar ratio, the predominant cleaved S2 band of SARS2-S was almost completely
88 diminished in whole cell lysate and in cell surface fraction (Fig.2a). Thus, F1 strongly interfered with
89 the expression and cell surface translocation of SARS2-S. In sharp contrast, F2 did not exhibit any
90 significant interference even at a tenfold molar ratio (Fig.2a).

91 Sequence comparison of 2002 SARS-S, 2012 MERS-S and COVID-19 SARS2-S suggested a
92 wide range of sequence identity levels among them (Fig.S2, S3). For instance, 2002 SARS-S shared
93 77% amino-acid sequence identity with full-length COVID-19 SARS2-S, while 2012 MERS-S was only
94 at 35% (Fig.S2, S3). If only considering the regions included in F1 polypeptide, the identical residues
95 became 94% for 2002 SARS-S and 42% for 2012 MERS-S (Fig.S2, S3). These spike proteins with a
96 wide range of sequence identity levels were ideal for testing the insensitivity of F1-induced interference
97 to amino-acid changes. Strikingly, even with F1-harboring plasmid at a twofold molar ratio, the cleaved
98 SARS-S and MERS-S S2 bands were almost completely diminished in whole cell lysate and in cell
99 surface fraction (Fig.2b,c). Thus, despite the low sequence identity, F1 also strongly interfered with the
100 expression and cell surface translocation of SARS-S and MERS-S glycoproteins. Consistent with the
101 robust interference activity exhibited by F1 (Fig.2), a high level of F1 polypeptide was constantly
102 detected in cell surface fraction (Fig.S4a-c). In marked contrast, a limited level of F2 polypeptide was
103 detected in cell surface fraction (Fig.S4a), in agreement with the non-interference of F2 (Fig.2a).

104 **F1 interfered with coronavirus spike proteins at the protein level**

105 In order to probe the mechanism of F1-mediated interference of coronavirus spike proteins, we
106 analyzed the mRNA levels of coronavirus spike proteins in the absence or presence of F1 or F2-carrying
107 plasmids (**Fig.S5, S6**). Clearly co-transfection of F1- or F2-harboring plasmids did not significantly
108 change the mRNA levels of coronavirus spike proteins, which were kept at a relatively constant level
109 (between 50%~150%) comparing to the samples that were transfected only with plasmids carrying the
110 respective coronavirus spike gene (**Fig.S5**). In sharp contrast, the mRNA levels of F1 were at the order
111 of $2^{6.5} \sim 2^{12}$ when normalized against endogenous GAPDH (**Fig.S6**), justifying the high potency of
112 F1-mediated interference. The mRNA levels of F2 were comparable with those of F1 (**Fig.S6a**), in
113 marked contrast with the low level of F2 polypeptide detected in cell surface fraction (**Fig.S4a**),
114 suggested that the non-interference of F2 (**Fig.2a**) was likely due to the instability of the synthesized F2
115 polypeptide. Collectively, these data suggested that F1-mediated interference of various coronavirus
116 spike glycoproteins was not at the mRNA level.

117 We next investigated whether SARS2-S and F1 directly interacted with each other to form non-
118 native oligomers in the cell. SARS2-S and F1 were each tagged with a monomeric green fluorescent
119 protein (GFP) variant at the extreme C-terminus, CFP for SARS2-S (termed as SARS2C) and YFP for
120 F1 (named as F1Y) (**Fig.S5d**). Since the expression of SARS2-S was completely diminished when F1-
121 and SARS2-S-carrying plasmids were co-transfected at a 2:1 molar ratio (**Fig.2a**), we used F1Y- and
122 SARS2C-containing plasmids at a reduced 1:1 molar ratio to transiently transfect HEK293T cells, and
123 monitored fluorescence resonance energy transfer (FRET) between them. The FRET ratio (FR) was
124 determined by using the three-cube approach²¹. Although the non-native oligomers formed by
125 SRARS2C and F1Y were expected to be highly unstable, FRET signals between them were robustly
126 detected (**Fig.S5e**), supporting direct interactions between SARS2-S and F1. Collectively, these results

127 suggested that F1-mediated interference of various coronavirus spike glycoproteins was at the protein
128 level.

129 **F1 minicircle potently interfered with expression and cell surface translocation of all three**
130 **coronavirus spike glycoproteins**

131 The high potency of F1 in interfering with expression and surface translocation of the spike
132 glycoproteins from coronaviruses that caused severe outbreaks or pandemic between 2002 to 2021
133 suggests that F1 has a high promise to become an effective therapeutic agent against different
134 coronavirus lineages over a long time period. Therefore, we sought to identify a convenient way to
135 deliver F1 for therapeutic purpose. The potent interference activity of F1 requires it to co-localize with
136 its target spike protein at the time of protein synthesis, therefore F1-encoding gene needs to be delivered
137 to the site of action. Minicircles are a type of newly developed DNA carriers for gene therapy²². The
138 main features of minicircles include the cleaner gene background with minimal viral or bacterial gene
139 elements, sustained high-level protein expression, and more importantly, the small size that may allow
140 the use of aerosols for drug delivery²³. The latter may be a distinct advantage against coronavirus-
141 caused respiratory diseases.

142 We made a F1 minicircle by inserting the F1 coding sequence into the parental minicircle
143 cloning vector pMC.CMV-MCS-SV40polyA (**Fig.3a**), and tested its efficacy in interfering with
144 expression and surface translocation of coronavirus spike glycoproteins. Compared to the controls
145 where no minicircle was used, the presence of F1 minicircles, at a merely 4.5-fold molar ratio, almost
146 completely abolished cell surface translocation of all three spike proteins (**Fig.3b-d**). It is important to
147 emphasis that this level of potent interference was achieved under the situation where pcDNA3.1-based
148 plasmids harboring coronavirus spike-coding genes were efficiently replicated in HEK293T cells, while

149 F1 minicircle cannot. Therefore, F1 minicircle demonstrated strong interference with the expression and
150 cell surface translocation of these coronavirus spike proteins.

151 **F1 minicircle reduced the level of SARS2-S protein on intact pseudoviruses and impaired** 152 **pseudovirus infectivity**

153 To investigate the consequences of the reduced cell surface translocation of coronavirus spike
154 proteins by F1 minicircle, we compared the level of SARS2-S protein on pseudoviruses generated using
155 luciferase-expressing, *env*-defective HIV-1 genome plasmid pRL4.3-Luc-R⁻E⁻ in the presence of
156 different molar ratios of control minicircle made from the empty parental vector (termed as MN501A) or
157 F1 minicircle. In order to make sure that only spike proteins anchored on the pseudovirus envelope
158 were accounted for, we employed QuickTiter Lentivirus Titer kit to precipitate intact pseudoviruses
159 from cleared supernatant prior to analysis by western blot. Impressively, even with a twofold molar
160 ratio of F1 minicircle, almost no SARS2-S was detected on the generated intact pseudoviruses (**Fig.4a**).
161 Not surprisingly, these pseudoviruses completely failed to infect hACE2-expressing HEK293T cells
162 (**Fig.4b**).

163 **Discussion**

164 In the current research, we present a series of compelling evidence that the F1 polypeptide
165 derived from COVID-19 SARS2-S potently reduced the expression and surface translocation of the
166 spike proteins from all three coronaviruses that caused regional outbreaks or global pandemic in the last
167 20 years, despite as low as 35% amino-acid sequence identity among them. Although extensive
168 mutations were found on SARS2-S proteins in recent SARS-CoV-2 variants³⁻⁵ (**Fig.1a**), the regions
169 corresponding to F1 polypeptide were highly conserved, as exemplified by the nearly identical sequence
170 of SARS-CoV-2 B.1.1.7-S with F1 (**Fig.S3b**). Therefore, F1 polypeptide is expected to be effective on

171 the spike proteins of almost any emerging COVID-19 SARS-CoV-2 variants in the future. Furthermore,
172 since the spike proteins of other human coronaviruses and 2012 MERS-CoV have a similar level of
173 sequence identity with F1 (**Fig.S3b**), we expect F1 to be equally effective on all these spike proteins.
174 Moreover, this polypeptide-based interference targeting the conserved sequences of coronavirus spike
175 protein to interfere with its functions represents a novel concept for therapeutics with high efficacy and
176 specificity, resistance to evolution and easy adaptability to other viral proteins in combating against
177 important viral pathogens.

178

179 **Figure captions**

180 **Figure 1. The concept of polypeptide-based protein interference against coronavirus spike**
181 **proteins. a).** Domain organization of COVID-19 SARS2-S, the mutations in recent variants and the
182 design of interfering polypeptides F1 and F2. SP: Signal peptide; NTD: N-terminal domain; RBD:
183 receptor-binding domain; SD1: subdomain 1; SD2: subdomain 2; FP: fusion peptide; HR1: heptad
184 repeat 1; HR2: heptad repeat 2; TM: transmembrane domain; CT: Cytoplasmic tail. The cleavage at
185 S1/S2 (red arrow) gives rise to N-terminal S1 fragment and C-terminal S2 fragment. The signal peptide
186 sequence at the extreme N-termini of F1 and F2 allowed the polypeptides to be translocated in the same
187 way as COVID-19 SARS2-S. At the extreme C-termini, SARS2-S had a C9 epitope recognized by C9-
188 rhodopsin antibody 1D4, while both F1 and F2 had a FLAG-tag. **b).** Diagram of polypeptide-based
189 interference targeting coronavirus spike proteins. Top row: in the normal situation, the spike proteins
190 were synthesized, folded and formed native spike oligomers, which were anchored on virion envelope.
191 Bottom row, interfering polypeptides formed non-native oligomers with the wild-type spike proteins,
192 thus reducing the level of native spike oligomers on the envelope of new virions.

193 **Figure 2. F1 significantly reduced expression and surface translocation of three coronavirus spike**
194 **glycoproteins. a).** COVID-19 SARS2-S. **b).** 2002 SARS-S. **c).** 2012 MERS-S. For panels **a-c)**, the
195 levels of S protein in whole cell lysate (left) or in cell surface fraction (right) were compared for
196 HEK293T cells transfected with S-containing plasmid only (Sample 1), or together with twofold
197 (Sample 2) or tenfold molar ratio (Sample 3) of F1-containing plasmid, and in panel **a)** with twofold
198 (Sample 4) or tenfold molar ratio (Sample 5) of F2-containing plasmid.

199 **Figure 3. F1 minicircle significantly reduced expression and surface translocation of three**
200 **coronavirus spike glycoproteins. a).** Diagram for the production of F1 minicircle used in this study.
201 **b).** COVID-19 SARS2-S; **c).** 2002 SARS-S; **d).** 2012 MERS-S. The levels of uncleaved S protein and
202 the cleaved S2 protein in whole cell lysate (left) or in cell surface fraction (right) were compared for
203 HEK293T cells transfected with S-containing plasmid only (Sample 1), or together with 4.5-fold
204 (Sample 2) or 22.5-fold molar ratio (Sample 3) of F1 minicircles.

205 **Figure 4. F1 minicircle significantly reduced the level of SARS2-S on intact pseudoviruses and**
206 **impaired pseudovirus infectivity. a).** The level of cleaved S2 protein was compared for intact
207 pseudovirus produced from HEK293T cells transfected with SARS2-S-containing plasmid and different
208 ratios of empty minicircle control, MN501A (Sample 1~4), or F1 minicircle (Sample 5~8). **b).**
209 Pseudovirus generated in the presence of F1 minicircle at a twofold molar ratio completely lost the
210 ability to infect hACE2-expressing HEK293T cells. The infectivity of pseudovirus generated in the
211 presence of MN501A minicircle at a twofold molar ratio was considered as 100%.

212 **Materials and Methods**

213 **Materials**

214 The pcDNA3.1 plasmids harboring the genes encoding SARS-CoV spike (GenBank accession
215 number AFR58740.1), SARS-CoV-2 spike (GenBank accession number QHD43416.1), and human
216 ACE2 (GenBank accession number NM_021804) were kind gifts from Dr. Fang Li¹⁴ (Addgene plasmid
217 No. 145031, 145032 and 145033, respectively). The pcDNA3.1 plasmids harboring the genes encoding
218 MERS-CoV spike (GenBank accession number QBM11748.1), F1 or F2 polypeptides were synthesized
219 by Genscript Biotech (Piscataway, NJ, USA). The parental minicircle vector pMC.CMV-MCS-
220 SV40polyA (Cat. No. MN501A-1), ZYCY10P3S2T *E. coli* minicircle producer strain²⁴ competent cells
221 (Cat. No. MN900A-1) and Arabinose Induction Solution (Cat. No. MN850A-1) were purchased from
222 System Biosciences (Palo Alto, CA, USA). The luciferase-expressing, *env*-defective HIV-1 genome
223 plasmid pRL4.3-Luc-R'E⁻ (Cat. No. 3418) was obtained through the NIH AIDS Reagent Program,
224 Division of AIDS, NIAID, NIH, USA. C9-rhodopsin antibody 1D4 HRP (Cat. No. sc57432 HRP) and
225 2',3'-cyclin nucleotide 3'-phosphodiesterase (CNPase) antibody (Cat. No. A01308) were purchased
226 from Santa Cruz Biotechnology, Inc. (Dallas, TX, USA) and Genscript Biotech (Piscataway, NJ, USA),
227 respectively. QuickTiter Lentivirus Titer kit (Cat. No. VPK-107) was purchased from Cell Biolabs Inc
228 (San Diego, CA, USA). Pierce Cell surface Protein Biotinylation and Isolation Kit (Cat. No. A44390)
229 was purchased from Thermo Scientific, Waltham, MA, USA. Quick-RNA miniprep Kit (Cat. No.
230 R1054) and ZymoPURE II Plasmid Maxiprep Kit (Cat. No. D4203) were obtained from Zymo
231 Research, Irvine, CA, USA. iScript Reverse Transcription Supermix (Cat. No. 1708840) was purchased
232 from Bio-rad, Hercules, CA, USA. Bimake SYBR Green qPCR Master Mix (Cat. No. B21203) was
233 obtained from Bimake, Houston, TX, USA. Lipofectine 3000 (Cat. No. L3000015) was obtained from
234 Invitrogen, Carlsbad, CA, USA. One-Glo EX Luciferase Assay System (Cat. No. E8110) was
235 purchased from Promega, Madison, WI, USA. HEK293T cells (Cat. No. CRL-11268) were purchased

236 from American Type Culture Collection (Manassas, VA, USA) and maintained in DMEM medium
237 containing 10% fetal bovine serum (FBS) at 37°C incubator supplied with 5% CO₂.

238 **Cell surface biotinylation and protein purification**

239 Cell surface biotinylation and protein purification were performed using Pierce Cell surface
240 Protein Biotinylation and Isolation Kit following the manufacturer's instruction. Briefly, cell surface
241 proteins on HEK293T cells were first labeled with Sulfo-NHS-SS-Biotin at 4°C for 30 minutes, which
242 were then stopped by adding Tris-buffered saline and further washed. After cells were lysed with Lysis
243 Buffer, lysate was cleared by centrifugation. Cleared lysate was incubated with NeutrAvidin Agarose to
244 allow binding of biotinylated proteins. After extensive wash, the bound proteins were eluted with
245 Elution Buffer containing 10 mM DTT. The cleared lysate ("Whole cell" fraction) and eluted proteins
246 ("Cell surface" fraction) were run on 10% SDS-PAGE and the spike proteins were detected by C9-
247 rhodopsin antibody 1D4 HRP. Endogenous membrane-anchored protein CNPase was detected by anti-
248 CNPase antibody and served as an internal control.

249 **Total RNA isolation and RT-qPCR**

250 Total RNA was purified using Quick-RNA miniprep Kit. Reverse transcription was carried out
251 using iScript Reverse Transcription Supermix. qPCR was performed using Bimake SYBR Green qPCR
252 Master Mix with the following primers:

253 1060 (SARS-S Forward): GTTCAAGGACGGCATCTACTT

254 1061 (SARS-S Reverse): ACGCTCTGGGACTTGTTATTC

255 1062 (SARS2-S Forward): GACAAAGTGCACCCTGAAGA

256 1063 (SARS2-S Reverse): GGGCACAGGTTGGTGATATT

257 1089 (MERS-S Forward): GAACGCCTCTCTGAACTCTTT

258 1090 (MERS-S Reverse): GTCCTCGGTGATGTTGTATGT

259 1091 (F1 and F2 Forward): GATTAGAGCCGCCGAGATTAG

260 1092 (F1 and F2 Reverse): GGACTGAGGGAAAGACATGAG

261 Since the synthesized genes of F1 and F2 were optimized for mammalian expression and
262 different from the gene coding for SARS2-S, the qPCR primers 1091 and 1092 were unique to F1 and
263 F2, while 1062 and 1063 were unique to SARS2-S.

264 **Minicircle production**

265 F1-coding gene was cloned into the minicircle parental vector pMC.CMV-MCS-SV40polyA
266 (MN501A) to yield MN501A-F1. MN501A or MN501A-F1 was transformed into ZYCY10P3S2T *E.*
267 *coli* minicircle producer strain competent cells following the manufacturer's instruction. The production
268 of minicircle DNA was induced with the addition of Arabinose Induction Solution. Minicircle DNA
269 was purified by using ZymoPURE II Plasmid Maxiprep Kit per the manufacturer's instruction.

270 **Pseudovirus generation, precipitation and concentration**

271 Pseudovirus generation followed the protocol reported earlier²⁵. Essentially, HEK293T cells
272 were seeded on 6-well plates the night before. The next day, pcDNA3.1-SARS2-S (0.6 µg) and pRL4.3-
273 Luc-R⁻E⁻ (0.6 µg) were used to transfect one-well HEK293T cells using Lipofectine 3000. MN501A
274 minicircle or F1 minicircle at indicated molar ratio was included in the transfection mixture. At 16
275 hours post-transfection, the HEK293T cells were fed with fresh medium. At 48 hours after medium
276 change, the supernatant of each well of the 6-well plates was harvested, and centrifuged at 300 g for 5
277 minutes to remove cell debris. Pseudoviruses were purified using QuickTiter Lentivirus Titer kit
278 following the manufacturer's instruction. The virus lysate was analyzed by western blot using
279 rhodopsin antibody 1D4 HRP for spike proteins, and FITC-conjugated anti-p24 mAb and HRP-

280 conjugated anti-FITC mAb for p24, which served as an internal control. A portion of pseudovirus-
281 containing supernatant was concentrated by PEG8000 and used for luciferase assay of cell entry.

282 **Luciferase assay of cell entry by pseudoviruses**

283 HEK293T cells were seeded on 100 mm dishes the night before. The next day, HEK293T cells
284 were transfected with 10 µg pcDNA3.1-hACE2 using Lipofectine 3000. At 16 hours post-transfection,
285 the cells were resuspended in FBS-free DMEM medium, and plated onto 96-well white plates to which
286 10 µL concentrated pseudoviruses was already added to each well. Two hours later, each well was
287 added with an equal volume of DMEM containing 20% FBS. The cells were further incubated for 36
288 hours, then an equal volume of One-Glo EX Luciferase Assay Reagent was added, after incubation for 3
289 minutes, the luminescence signals were recorded.

290 **FRET between SARS2C and F1Y**

291 High quality/high resolution automated imaging was performed on a GE Healthcare DVLIVE
292 epifluorescence image restoration microscope using an Olympus PlanApoN 60X/1.42 NA objective and
293 a 1.9k x 1.9k pco.EDGE sCMOS_5.5 camera with a 1024x1024 FOV. The filter sets used were: CFP
294 (438/24 excitation, 470/24 emission) and YFP (513/17 excitation, 559/38 emission). Donor and
295 Acceptor control channels were acquired using CFP-CFP and YFP-YFP, respectively. FRET images
296 were acquired using CFP-YFP excitation and emission filter pair. Cells were chosen with similar
297 intensity profiles in donor and acceptor prior to acquisition while under 37°C and 5% CO₂
298 environmental conditions. Z stacks (0.25µm) covering the whole cell (~12µm) were acquired before
299 applying a conservative restorative algorithm for quantitative image deconvolution using SoftWorx
300 v7.0, and saving files as a max pixel intensity projection tiff for each individual channel. The FRET

ratio (FR) was determined by using the three-cube approach²¹ that was defined as the ratio of YFP emission in the presence of FRET over that in the absence of FRET:

$$FR = F_{A(D)} / F_A = [S_{FRET(DA)} - R_{D1} \cdot S_{CFP(DA)}] / R_{A1} \cdot [S_{YFP(DA)} - R_{D2} \cdot S_{CFP(DA)}]$$

Where $R_{D1} = S_{FRET(D)} / S_{CFP(D)}$; $R_{D2} = S_{YFP(D)} / S_{CFP(D)}$; $R_{A1} = S_{FRET(A)} / S_{YFP(A)}$; and S_{FRET} , S_{CFP} and S_{YFP} with D, A, DA in parenthesis refer to fluorescence signals of the FRET, CFP or YFP channel when only SARS2C (containing CFP as donor (D)), F1Y (containing YFP as acceptor (A)), or both SARS2C and F1Y-containing plasmids (both donor and acceptor (DA)) were included in the transfection. Averages from untransfected cells were subtracted from same-day fluorescence values for each filter cube. According to the definition, $FR = 1.0$ means no FRET, while $FR > 1.0$ indicates the presence of FRET between the proteins carrying donor and acceptor proteins.

References

- 1 Cui, J., Li, F. & Shi, Z. L. Origin and evolution of pathogenic coronaviruses. *Nature reviews. Microbiology* **17**, 181-192, doi:10.1038/s41579-018-0118-9 (2019).
- 2 de Wit, E., van Doremalen, N., Falzarano, D. & Munster, V. J. SARS and MERS: recent insights into emerging coronaviruses. *Nature reviews. Microbiology* **14**, 523-534, doi:10.1038/nrmicro.2016.81 (2016).
- 3 Rambaut, A. *et al.* Preliminary genomic characterisation of an emergent SARS-CoV-2 lineage in the UK defined by a novel set of spike mutations. (2020).
- 4 Tegally, H., Wilkinson, E. & al., e. Emergence and rapid spread of a new severe acute respiratory syndrome-related coronavirus 2 (SARS-CoV-2) lineage with multiple spike mutations in South Africa. *medRxiv* **2020.2012.2021.20248640** (2020).
- 5 Naveca, F. *et al.* Phylogenetic relationship of SARS-CoV-2 sequences from Amazonas with emerging Brazilian variants harboring mutations E484K and N501Y in the Spike protein. *virological.org* (2021).
- 6 Zhou, D., Dejnirattisai, W., Supasa, P., Liu, C. & al., e. Evidence of escape of SARS-CoV-2 variant B.1.351 from natural and vaccine induced sera. *Cell* **In press** (2021).
- 7 Tada, T. *et al.* Neutralization of viruses with European, South African, and United States SARS-CoV-2 variant spike proteins by convalescent sera and BNT162b2 mRNA vaccine-elicited antibodies. *bioRxiv*, doi:10.1101/2021.02.05.430003 (2021).
- 8 Wang, P. *et al.* Increased Resistance of SARS-CoV-2 Variants B.1.351 and B.1.1.7 to Antibody Neutralization. *bioRxiv*, doi:10.1101/2021.01.25.428137 (2021).

- 332 9 Zhou, B. *et al.* SARS-CoV-2 spike D614G change enhances replication and transmission.
333 *Nature*, doi:10.1038/s41586-021-03361-1 (2021).
- 334 10 Cele, S. *et al.* Escape of SARS-CoV-2 501Y.V2 from neutralization by convalescent plasma.
335 *Nature*, doi:10.1038/s41586-021-03471-w (2021).
- 336 11 V'Kovski, P., Kratzel, A., Steiner, S., Stalder, H. & Thiel, V. Coronavirus biology and
337 replication: implications for SARS-CoV-2. *Nature reviews. Microbiology* **19**, 155-170,
338 doi:10.1038/s41579-020-00468-6 (2021).
- 339 12 Wrapp, D. *et al.* Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation.
340 *Science* **367**, 1260-1263, doi:10.1126/science.abb2507 (2020).
- 341 13 Walls, A. C. *et al.* Structure, Function, and Antigenicity of the SARS-CoV-2 Spike
342 Glycoprotein. *Cell* **183**, 1735, doi:10.1016/j.cell.2020.11.032 (2020).
- 343 14 Shang, J. *et al.* Structural basis of receptor recognition by SARS-CoV-2. *Nature* **581**, 221-224,
344 doi:10.1038/s41586-020-2179-y (2020).
- 345 15 Cai, Y. *et al.* Distinct conformational states of SARS-CoV-2 spike protein. *Science* **369**, 1586-
346 1592, doi:10.1126/science.abd4251 (2020).
- 347 16 Bangaru, S. *et al.* Structural analysis of full-length SARS-CoV-2 spike protein from an advanced
348 vaccine candidate. *Science* **370**, 1089-1094, doi:10.1126/science.abe1502 (2020).
- 349 17 Klein, S. *et al.* SARS-CoV-2 structure and replication characterized by in situ cryo-electron
350 tomography. *Nat Commun* **11**, 5885, doi:10.1038/s41467-020-19619-7 (2020).
- 351 18 Stertz, S. *et al.* The intracellular sites of early replication and budding of SARS-coronavirus.
352 *Virology* **361**, 304-315, doi:10.1016/j.virol.2006.11.027 (2007).
- 353 19 Coutard, B. *et al.* The spike glycoprotein of the new coronavirus 2019-nCoV contains a furin-
354 like cleavage site absent in CoV of the same clade. *Antiviral Res* **176**, 104742,
355 doi:10.1016/j.antiviral.2020.104742 (2020).
- 356 20 Johnson, B. A. *et al.* Loss of furin cleavage site attenuates SARS-CoV-2 pathogenesis. *Nature*
357 **591**, 293-299, doi:10.1038/s41586-021-03237-4 (2021).
- 358 21 Erickson, M. G., Alseikhan, B. A., Peterson, B. Z. & Yue, D. T. Preassociation of calmodulin
359 with voltage-gated Ca(2+) channels revealed by FRET in single living cells. *Neuron* **31**, 973-
360 985, doi:10.1016/s0896-6273(01)00438-x (2001).
- 361 22 Hardee, C. L., Arevalo-Soliz, L. M., Hornstein, B. D. & Zechiedrich, L. Advances in Non-Viral
362 DNA Vectors for Gene Therapy. *Genes (Basel)* **8**, doi:10.3390/genes8020065 (2017).
- 363 23 Catanese, D. J., Jr., Fogg, J. M., Schrock, D. E., 2nd, Gilbert, B. E. & Zechiedrich, L.
364 Supercoiled Minivector DNA resists shear forces associated with gene therapy delivery. *Gene*
365 *Ther* **19**, 94-100, doi:10.1038/gt.2011.77 (2012).
- 366 24 Kay, M. A., He, C. Y. & Chen, Z. Y. A robust system for production of minicircle DNA vectors.
367 *Nat Biotechnol* **28**, 1287-1289, doi:10.1038/nbt.1708 (2010).
- 368 25 Zhao, G. *et al.* A safe and convenient pseudovirus-based inhibition assay to detect neutralizing
369 antibodies and screen for viral entry inhibitors against the novel human coronavirus MERS-CoV.
370 *Virol J* **10**, 266, doi:10.1186/1743-422X-10-266 (2013).
- 371 26 Madeira, F. *et al.* The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic*
372 *Acids Res* **47**, W636-W641, doi:10.1093/nar/gkz268 (2019).

373

374 Supplementary Information

375 **Figure S1. Amino-acid sequences of F1 and F2 polypeptides used in this study.** Single-letter codes
376 were used for amino acids. The signal peptide sequences were highlighted in boldface at the N-
377 terminus. FLAG-tags at the extreme C-terminus of each polypeptide were underlined.

378 **Figure S2. Sequence alignments of coronavirus spike proteins. a).** The alignment of SARS2-S
379 (SARS2), SARS-S (SARS), and SARS-CoV-2 B.1.1.7 spike protein (B.1.1.7). **b).** The alignment of
380 SARS2-S (SARS2) and MERS-S (MERS) proteins. In both panels, the range of residues included in the
381 F1 polypeptide was highlighted in blue boxes. The cleavage site between S1 and S2 was shown by a red
382 arrow. Multiple sequence alignments were performed using CLUSTAL Omega²⁶. The sequences were:
383 SARS2-S (GenBank accession number AFR58740.1), SARS-S (GenBank accession number
384 QHD43416.1), SARS-CoV-2 B.1.1.7-S (GenBank accession number QQH18545.1), and MERS-S
385 (GenBank accession number QBM11748.1).

386 **Figure S3. Sequence conservation of representative human coronavirus spike proteins comparing**
387 **to SARS2-S. a).** The comparison of SARS-S, B.1.1.7-S and MERS-S proteins with full-length SARS2-
388 S. **b).** The comparison of representative human coronavirus spike proteins with SARS2-S F1
389 polypeptide (residues 911~1273). In both panels, shown were the number of residues that were identical
390 (Identity), similar (Positive) and unmatched (Gap) comparing to SARS2-S, the total residues compared
391 (after “/”) and the corresponding percentage (in parenthesis). The total residues varied due to gaps in
392 sequence alignments. The sequences were: SARS2-S (GenBank accession number AFR58740.1),
393 SARS-S (GenBank accession number QHD43416.1), SARS-CoV-2 B.1.1.7-S (GenBank accession
394 number QQH18545.1), MERS-S (GenBank accession number QBM11748.1), HCoV-HKU1-S (Uniprot
395 accession number Q0ZME7), HCoV-OC43-S (Uniprot accession number P36334), HCoV-NL63-S
396 (Uniprot accession number Q6Q1S2), and HCoV-229E-S (Uniprot accession number P15423).

397 **Figure S4. Protein expression levels of F1 and F2 when co-transfected with coronavirus spike**
398 **protein-coding plasmids. a).** with COVID-19 SARS2-S; **b).** with 2002 SARS-S; **c).** with 2012 MERS-
399 S.

400 **Figure S5. F1 interfered with coronavirus spike proteins at the protein level. a~c).** The mRNA
401 levels of spike proteins in the absence or presence of F1-carrying plasmid. **a).** COVID-19 SARS2-S; **b).**
402 2002 SARS-S; **c).** 2012 MERS-S. **d).** Diagram for the FRET setup, where CFP and YFP were each
403 tagged to the extreme C-terminus of SARS2-S and F1, resulting in SARS2C and F1Y, respectively. **e).**
404 FRET ratio of SARS2C with F1Y (at a 1:1 molar ratio for transfection), compared to that of F1Y only.
405 ****, $p < 0.0001$ in unpaired *t*-test comparing FRET ratios of SARS2C+F1Y *versus* same-day F1Y only.
406 Numbers in parentheses indicate the number of cells included in each analysis.

407 **Figure S6. The mRNA levels of F1 and F2 when co-transfected with coronavirus spike protein-**
408 **coding plasmids. a).** with COVID-19 SARS2-S; **b).** with 2002 SARS-S; **c).** with 2012 MERS-S.

409

Figure 1. The concept of polypeptide-based protein interference against coronavirus spike proteins. **a).** Domain organization of COVID-19 SARS2-S, the mutations in recent variants and the design of interfering polypeptides F1 and F2. SP: Signal peptide; NTD: N-terminal domain; RBD: receptor-binding domain; SD1: subdomain 1; SD2: subdomain 2; FP: fusion peptide; HR1: heptad repeat 1; HR2: heptad repeat 2; TM: transmembrane domain; CT: Cytoplasmic tail. The cleavage at S1/S2 (red arrow) gives rise to N-terminal S1 fragment and C-terminal S2 fragment. The signal peptide sequence at the extreme N-termini of F1 and F2 allowed the polypeptides to be translocated in the same way as COVID-19 SARS2-S. At the extreme C-termini, SARS2-S had a C9 epitope recognized by C9-rhodopsin antibody 1D4, while both F1 and F2 had a FLAG-tag. **b).** Diagram of polypeptide-based interference targeting coronavirus spike proteins. Top row: in the normal situation, the spike proteins were synthesized, folded and formed native spike oligomers, which were anchored on virion envelope. Bottom row, interfering polypeptides formed non-native oligomers with the wild-type spike proteins, thus reducing the level of native spike oligomers on the envelope of new virions.

Figure 2. F1 significantly reduced expression and surface translocation of three coronavirus spike glycoproteins. a). COVID-19 SARS2-S. b). 2002 SARS-S. c). 2012 MERS-S. For panels a-c), the levels of S protein in whole-cell lysate (left) or in cell-surface fraction (right) were compared for HEK293T cells transfected with S-containing plasmid only (Sample 1), or together with twofold (Sample 2) or tenfold molar ratio (Sample 3) of F1-containing plasmid, and in panel a) with twofold (Sample 4) or tenfold molar ratio (Sample 5) of F2-containing plasmid.

Figure 3. F1 minicircle significantly reduced expression and surface translocation of three coronavirus spike glycoproteins. **a).** Diagram for the production of F1 minicircle used in this study. **b).** COVID-19 SARS2-S; **c).** 2002 SARS-S; **d).** 2012 MERS-S. The levels of uncleaved S protein and the cleaved S2 protein in whole-cell lysate (left) or in cell-surface fraction (right) were compared for HEK293T cells transfected with S-containing plasmid only (Sample 1), or together with 4.5-fold (Sample 2) or 22.5-fold molar ratio (Sample 3) of F1 minicircles.

Figure 4. F1 minicircle significantly reduced the level of SARS2-S on intact pseudoviruses and impaired pseudovirus infectivity. **a).** The level of cleaved S2 protein was compared for intact pseudovirus produced from HEK293T cells transfected with SARS2-S-containing plasmid and different ratios of empty minicircle control, MN501A (Sample 1~4), or F1 minicircle (Sample 5~8). **b).** Pseudovirus generated in the presence of only twofold molar ratio of F1 minicircle completely lost the ability to infect hACE2-expressing HEK293T cells. The infectivity of pseudovirus generated in the presence of twofold molar ratio of MN501A minicircle was considered as 100%.

F1:

MFVFLVLLPLVSSQCVNLTRTQVTQNVLYENQKLIANQFNSAIGKIQDSLSTASALGKLQDVVNQNAQ
ALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAAT
KMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHGKAHFPREGVFVS
NGTHWFVTQRNFYEPQIIITDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLG
DISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGLIAIVMTIMLCCM
TSCCCLKGCCSCGSCCKFDEDDSEPVLGVKLHYTDYKDDDDK

F2:

MFVFLVLLPLVSSQCVNLTRTQDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSE
CVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHGKAHFPREGVFVSNGTH
WFVTQRNFYEPQIIITDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDIG
INASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGLIAIVMTIMLCCMTSCC
SCLKGCCSCGSCCKFDEDDSEPVLGVKLHYTDYKDDDDK

Figure S1. Amino-acid sequences of F1 and F2 polypeptides used in this study. Single-letter codes were used for amino acids. The signal peptide sequences were highlighted in boldface at the N-terminus. FLAG-tags at the extreme C-terminus of each polypeptide were underlined.

a)

```

SARS      MPMGSLQPLATLYLLGMLVASVLSGSDLDRCCTTFDDVQAPNYTQHTSSMRGVYYPDEIF  SDTLYLTQDLFLFFYSNVTGFHTIN-----HTFGNPFVIFPKDGIYFAATEKSNVVRG  113
SARS2     -----MFVFLVLLPLV-----SSQCVMLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLFFSNVTWFHAIHVSQGTNGTKRFDNPVLPFNDGVYFASTEKSNIIIRG  103
B.1.1.7   -----MFVFLVLLPLV-----SSQCVMLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLFFSNVTWFHAI--SGTNGTKRFDNPVLPFNDGVYFASTEKSNIIIRG  101
          :::: :*: *          :.*.:          :.* *****:.....*: *          :.*.:**:*:*:*:*:*:*:*:*:*

SARS      WVFSGTMNKSQSVIIINNSTNVIRACNFCNPFPAVSKPMGT---QTHMIFDNFACNTEFYISDAFSLDVSEKSGNFKHLREFVFNKNDGFLYVYKGYQPIDVVRDLPSGFNTL  229
SARS2     WIFGTTLDSTQSLIVNNATNVIVKCFQFCNDPFLGVYVYHNNKSWMESEFRVYSSANNCTFEYVSQPFMDLEGRQGNFKNLRFPVFNKIDGYFKIYKHTPINLVRDLPPQGSAL  223
B.1.1.7   WIFGTTLDSTQSLIVNNATNVIVKCFQFCNDPFLGVY-HKNNKSWMESEFRVYSSANNCTFEYVSQPFMDLEGRQGNFKNLRFPVFNKIDGYFKIYKHTPINLVRDLPPQGSAL  220
          *:***:..*:**:*:*:*:*:*:*:*:*:*:*:*..  ..  :..  :..* *****: * :*.. *.****:*:*:*:* **: :. : *:*:*:*:*..*:*

SARS      KPIFKPLPGINITRFRALTAFS-----PAQDIWGTSAAYFVGYLKPTTFLMKYDENGITDAVDCSQNPLAELKCSVKSFEDKGIYQTSNFRVVPVSGDVVRFFNITNLCPFGEVFN  343
SARS2     EPLVDLPIGINITRFTQLLALHRSYLTGDDSSSGWTAGAAAYVGYLQPTFTLLKYNENGITDAVDCALDPLSEKTKLKSFTVEKGIYQTSNFRVVPVQTESIVRFFNITNLCPFGEVFN  343
B.1.1.7   EPLVDLPIGINITRFTQLLALHRSYLTGDDSSSGWTAGAAAYVGYLQPTFTLLKYNENGITDAVDCALDPLSEKTKLKSFTVEKGIYQTSNFRVVPVQTESIVRFFNITNLCPFGEVFN  340
          :*..**:*:*:*:*:*:*:*..  ..  * :.*:*:*:*:* * *:*:*:*:*:*:*:*:*:* * :.*:*:*:*:* * :.*:*:*:*:*

SARS      ATKFPVYAWERKKISNCVADSVLYNSTFFSTFKCYGVSAATKLNLCFSNVYADSVFVKGDDVRQIAPGQTGVADYNYKLPDDFMGCVLAWNTRNIDATSTGNYNKYRYLRHKGKLRP  463
SARS2     ATRFASVYAWNRRKISNCVADSVLYNSASFSTFKCYGVSPTKLNLCFTNVYADSVFVIRGDEVQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNLDKSVGGNYNYLYRFRKSNLKP  463
B.1.1.7   ATRFASVYAWNRRKISNCVADSVLYNSASFSTFKCYGVSPTKLNLCFTNVYADSVFVIRGDEVQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNLDKSVGGNYNYLYRFRKSNLKP  460
          **:* *****:.*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*

SARS      FERDISNVFPSPDGKPCPTP-PALNCYWPIMDYGFTTGGIGYQPYRVVVLSPFELNAPATVCGPKLSTDLIKNQCVNFNENGLTGTGVLTPSKRFPFQQFGRDVSDFPTSDVRDPKTS  582
SARS2     FERDISTEIQAGSTPCNGVEGFNCYFPLQSYGFPQTNGVGYQPYRVVVLSPFELHAPATVCGPKKSTNLVKNKCVNFNENGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTL  583
B.1.1.7   FERDISTEIQAGSTPCNGVEGFNCYFPLQSYGFPQTNGVGYQPYRVVVLSPFELHAPATVCGPKKSTNLVKNKCVNFNENGLTGTGVLTESNKKFLPFQQFGRDIDTTDAVRDPQTL  580
          *****  :.  :..*..  :.*:*:*:*:* * *:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:* * *:*:*:*:* *

SARS      ILDISPCSPGGVSVITPGTNTASSEVAVLYQDVNCTDVSTAIHADQLTPAWRIYSTGNVVFQTQAGCLIGAEHVDTSEYCDIPIGAGICASYHTVSL---LRSTSQKSIVAYTMSLGADS  698
SARS2     ILDITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTRVYSTGNSVVFQTRAGCLIGAEHVNSYECDIPIGAGICASYQTQTSNRRARSVASQSIIAYTMSLGAEN  703
B.1.1.7   ILDITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTRVYSTGNSVVFQTRAGCLIGAEHVNSYECDIPIGAGICASYQTQTSNRRARSVASQSIIAYTMSLGAEN  700
          *****:.*:*:*:*:*:*:*:* * *:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:* * :.*:*:*:*:*:*:*:*..

SARS      SIAYSNTTIAIPTNISISITTEVMPVSMAKTSVDCNMVYICGDSSTECANLLQYGSFCTQLNRLSGIAAEQDRNTRREVFAQVKQMYKTPTLKYFGGFNFSQILPDPLKPTKRSFIEDLLF  818
SARS2     SVAYSNNISIAIPTNFTISVTTEILPVSMKTTSVDCNTMYICGDSSTECANLLQYGSFCTQLNRLSGIAAEQDRNTRREVFAQVKQMYKTPTLKYFGGFNFSQILPDPLKPTKRSFIEDLLF  823
B.1.1.7   SVAYSNNISIAIPTNFTISVTTEILPVSMKTTSVDCNTMYICGDSSTECANLLQYGSFCTQLNRLSGIAAEQDRNTRREVFAQVKQMYKTPTLKYFGGFNFSQILPDPLKPTKRSFIEDLLF  820
          *:*****:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*

SARS      NKVTLADAGFMKQYGECLGINARDLICAKFNGLTVLPPLTTDDMIAAYTAALVSGTATAGWTFGAAGALQIPFAMQAMAYRFGNIIVTQNVLYENQKLIANQFNKAIISQIESLTTTST  938
SARS2     NKVTLADAGFMKQYGECLGINARDLICAKFNGLTVLPPLTTDDMIAAYTAALVSGTATAGWTFGAAGALQIPFAMQAMAYRFGNIIVTQNVLYENQKLIANQFNKAIISQIESLTTTST  943
B.1.1.7   NKVTLADAGFMKQYGECLGINARDLICAKFNGLTVLPPLTTDDMIAAYTAALVSGTATAGWTFGAAGALQIPFAMQAMAYRFGNIIVTQNVLYENQKLIANQFNKAIISQIESLTTTST  940
          *****:.*:*:*:*:*:*:* * *:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:* * *:*:*:*:*:*:*:*:*:*:*:*:*

SARS      ALGKLQDVVNQNAQALNTLVKQLSSNFGAISVVLNDILSRLDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMESECVLGQSKRVDFCGKGYHLSMFPQAPHGTVFL  1058
SARS2     ALGKLQDVVNQNAQALNTLVKQLSSNFGAISVVLNDILSRLDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMESECVLGQSKRVDFCGKGYHLSMFPQAPHGTVFL  1063
B.1.1.7   ALGKLQDVVNQNAQALNTLVKQLSSNFGAISVVLNDILARLDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMESECVLGQSKRVDFCGKGYHLSMFPQAPHGTVFL  1060
          *****:.*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*

SARS      HVTYVPSQERNFTTAPAIQCHGKAHFFPREGVVFVNGTTSWFITQRNFSPQIITTDNTFVSGNCDVIGIINNNTVYDPLQPELDSFKBELDKYFRNHTSPDVLGDISGINASVNNIQKEI  1178
SARS2     HVTYVPAQEKNTTAPAIQCHDGAHFFPREGVVFVNGTTSWFITQRNFYEPQIITTDNTFVSGNCDVIGIINNNTVYDPLQPELDSFKBELDKYFRNHTSPDVLGDISGINASVNNIQKEI  1183
B.1.1.7   HVTYVPAQEKNTTAPAIQCHDGAHFFPREGVVFVNGTTSWFITQRNFYEPQIITTDNTFVSGNCDVIGIINNNTVYDPLQPELDSFKBELDKYFRNHTSPDVLGDISGINASVNNIQKEI  1180
          *****:.*:*:*:*:*:*:* * *:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:* * *:*:*:*:*:*:*:*:*:*:*:*

SARS      DRLNEVAKNLNESLIDLQELGKYEYIKWPYWIWLGFIAGLIAIVMTILLCCMTSCCSCCLGACSCGSCCKFDEDDSEPVLGKVKLYH  1268
SARS2     DRLNEVAKNLNESLIDLQELGKYEYIKWPYWIWLGFIAGLIAIVMTILLCCMTSCCSCCLGKCCSCGSCCKFDEDDSEPVLGKVKLYH  1273
B.1.1.7   DRLNEVAKNLNESLIDLQELGKYEYIKWPYWIWLGFIAGLIAIVMTILLCCMTSCCSCCLGKCCSCGSCCKFDEDDSEPVLGKVKLYH  1270
          *****:.*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*

```

Figure S2. Sequence alignments of coronavirus spike proteins. a). The alignment of SARS2-S (SARS2), SARS-S (SARS), and SARS-CoV-2 B.1.1.7 spike protein (B.1.1.7). b). The alignment of SARS2-S (SARS2) and MERS-S (MERS) proteins. In both panels, the range of residues included in the F1 polypeptide was highlighted in blue boxes. The cleavage site between S1 and S2 was shown by a red arrow. Multiple sequence alignments were performed using CLUSTAL Omega²⁴. The sequences were: SARS2-S (GenBank accession number AFR58740.1), SARS-S (GenBank accession number QHD43416.1), SARS-CoV-2 B.1.1.7-S (GenBank accession number QQH18545.1), and MERS-S (GenBank accession number QBM11748.1).

Figure S2. Sequence alignments of coronavirus spike proteins. a) The alignment of SARS2-S (SARS2), SARS-S (SARS), and SARS-CoV-2 B.1.1.7 spike protein (B.1.1.7). **b)** The alignment of SARS2-S (SARS2) and MERS-S (MERS) proteins. In both panels, the range of residues included in the F1 polypeptide was highlighted in blue boxes. The cleavage site between S1 and S2 was shown by a red arrow. Multiple sequence alignments were performed using CLUSTAL Omega²⁴. The sequences were: SARS2-S (GenBank accession number AFR58740.1), SARS-S (GenBank accession number QHD43416.1), SARS-CoV-2 B.1.1.7-S (GenBank accession number QQH18545.1), and MERS-S (GenBank accession number QBM11748.1).

a)

	With SARS2-S (1~1273)		
Spike protein	Identity	Positive	Gap
SARS-S	963/1250 (77%)	1097/1250 (87%)	24/1250 (1%)
B.1.1.7-S	1263/1273 (99%)	1264/1273 (99%)	3/1273 (0%)
MERS-S	370/1057 (35%)	545/1057 (51%)	76/1057 (7%)

b)

	With SARS2-S F1 (911~1273)		
Spike protein	Identity	Positive	Gap
SARS-S (<i>Betacoronavirus</i> , lineage B)	340/363 (94%)	356/363 (98%)	0/363 (0%)
B.1.1.7-S (<i>Betacoronavirus</i> , lineage B)	361/363 (99%)	362/363 (99%)	0/363 (0%)
MERS-S (<i>Betacoronavirus</i> , lineage C)	156/368 (42%)	219/368 (59%)	20/363 (5%)
HCoV-HKU1-S (<i>Betacoronavirus</i> , lineage A)	154/347 (44%)	218/347 (62%)	12/347 (3%)
HCoV-OC43-S (<i>Betacoronavirus</i> , lineage A)	160/358 (45%)	224/358 (62%)	6/358 (1%)
HCoV-NL63-S (<i>Alphacoronavirus</i>)	136/386 (35%)	202/386 (52%)	56/386 (14%)
HCoV-229E-S (<i>Alphacoronavirus</i>)	135/377 (36%)	201/377 (53%)	46/377 (12%)

Figure S3. Sequence conservation of representative human and bat coronavirus spike proteins comparing to SARS2-S. **a).** The comparison of SARS-S, B.1.1.7-S and MERS-S proteins with full-length SARS2-S. **b).** The comparison of representative human and bat coronavirus spike proteins with SARS2-S F1 polypeptide (residues 911~1273). In both panels, shown were the number of residues that were identical (Identity), similar (Positive) and unmatched (Gap) comparing to SARS2-S, the total residues compared (after “/”) and the corresponding percentage (in parenthesis). The total residues varied due to gaps during sequence alignments. The sequences were: SARS2-S (GenBank accession number AFR58740.1), SARS-S (GenBank accession number QHD43416.1), SARS-CoV-2 B.1.1.7-S (GenBank accession number QQH18545.1), MERS-S (GenBank accession number QBM11748.1), HCoV-HKU1-S (Uniprot accession number Q0ZME7), HCoV-OC43-S (Uniprot accession number P36334), HCoV-NL63-S (Uniprot accession number Q6Q1S2), and HCoV-229E-S (Uniprot accession number P15423).

Figure S4. Protein expression levels of F1 and F2 when co-transfected with coronavirus spike protein-coding plasmids. a). with COVID-19 SARS2-S; b). with 2002 SARS-S; c). with 2012 MERS-S.

Figure S5. F1 interfered with coronavirus spike proteins at the protein level. a-c). The mRNA levels of spike proteins in the absence or presence of F1-carrying plasmid. a). COVID-19 SARS2-S; b). 2002 SARS-S; c). 2012 MERS-S. d). Diagram for the FRET setup, where CFP and YFP were each tagged to the extreme C-terminus of SARS2-S and F1, resulting in SARS2C and F1Y, respectively. e). FRET ratio of SARS2C with F1Y (at a 1:1 molar ratio for transfection), compared to that of F1Y only. ****, $p < 0.0001$ in unpaired t -test comparing FRET ratios of SARS2C+F1Y versus same-day F1Y only. Numbers in parentheses indicate the number of cells included in each analysis.

Figure S6. The mRNA levels of F1 and F2 when co-transfected with coronavirus spike protein-coding plasmids. a). with COVID-19 SARS2-S; b). with 2002 SARS-S; c). with 2012 MERS-S.

Figures

Figure 1

The concept of polypeptide-based protein interference against coronavirus spike proteins. a). Domain organization of COVID-19 SARS2-S, the mutations in recent variants and the design of interfering polypeptides F1 and F2. SP: Signal peptide; NTD: N-terminal domain; RBD: receptor-binding domain; SD1: subdomain 1; SD2: subdomain 2; FP: fusion peptide; HR1: heptad repeat 1; HR2: heptad repeat 2; TM: transmembrane domain; CT: Cytoplasmic tail. The cleavage at S1/S2 (red arrow) gives rise to N-terminal S1 fragment and C-terminal S2 fragment. The signal peptide sequence at the extreme N-termini of F1 and F2 allowed the polypeptides to be translocated in the same way as COVID-19 SARS2-S. At the extreme C-termini, SARS2-S had a C9 epitope recognized by C9-rhodopsin antibody 1D4, while both F1 and F2 had a FLAG-tag. b). Diagram of polypeptide-based interference targeting coronavirus spike proteins. Top row: in the normal situation, the spike proteins were synthesized, folded and formed native spike oligomers, which were anchored on virion envelope. Bottom row, interfering polypeptides formed non-native

oligomers with the wild-type spike proteins, thus reducing the level of native spike oligomers on the envelope of new virions.

Figure 2

F1 significantly reduced expression and surface translocation of three coronavirus spike glycoproteins. a). COVID-19 SARS2-S. b). 2002 SARS-S. c). 2012 MERS-S. For panels a-c), the levels of S protein in whole-cell lysate (left) or in cell-surface fraction (right) were compared for HEK293T cells transfected with S-containing plasmid only (Sample 1), or together with twofold (Sample 2) or tenfold molar ratio (Sample 3) of F1-containing plasmid, and in panel a) with twofold (Sample 4) or tenfold molar ratio (Sample 5) of F2-containing plasmid.

Figure 3

F1 minicircle significantly reduced expression and surface translocation of three coronavirus spike glycoproteins. a). Diagram for the production of F1 minicircle used in this study. b). COVID-19 SARS2-S; c). 2002 SARS-S; d). 2012 MERS-S. The levels of uncleaved S protein and the cleaved S2 protein in whole-cell lysate (left) or in cell-surface fraction (right) were compared for HEK293T cells transfected with S-containing plasmid only (Sample 1), or together with 4.5-fold (Sample 2) or 22.5-fold molar ratio (Sample 3) of F1 minicircles.

Figure 4

F1 minicircle significantly reduced the level of SARS2-S on intact pseudoviruses and impaired pseudovirus infectivity. a). The level of cleaved S2 protein was compared for intact pseudovirus produced from HEK293T cells transfected with SARS2-S-containing plasmid and different ratios of empty minicircle control, MN501A (Sample 1~4), or F1 minicircle (Sample 5~8). b). Pseudovirus generated in the presence of only twofold molar ratio of F1 minicircle completely lost the ability to infect hACE2-expressing HEK293T cells. The infectivity of pseudovirus generated in the presence of twofold molar ratio of MN501A minicircle was considered as 100%.