1 Identification of the Fusion Peptide-Containing Region in Betacoronavirus Spike Glycoproteins

2 Xiuyuan Ou[1&], Wangliang Zheng[1&], Yiwei Shan[1], Zhixia Mu[1], Samuel R. Dominguez[2], Kathryn
3 V. Holmes[3], and Zhaohui Qian[1]
4
5 MOH Key laboratory of Systems Biology of Pathogens, Institute of Pathogen Biology, Chinese
6 Academy of Medical Sciences and Peking Union Medical College[1], Beijing, 100176, China;
7 Department of Pediatrics[2], Department of Microbiology[3], University of Colorado School of
8 Medicine, Aurora, CO 80045
9
10 Key words: coronavirus spike glycoprotein, coronavirus fusion peptide, coronavirus membrane
11 fusion, MERS-CoV entry, SARS-CoV, MHV
12 Running title: Fusion peptide of spike protein of betacoronavirus
13
14 & XO and WZ contributed equally to this study.
15 #Address correspondence to Zhaohui Qian, zqian2013@sina.com,
16
17
18

Abstract (250 words)

The fusion peptides (FP) play an essential role in fusion of viral envelope with cellular membranes. The location and properties of the FPs in the spike (S) glycoproteins of different coronaviruses (CoV) have not yet been determined. Through amino acid sequence analysis of S proteins of representative CoVs, we identified a common region as a possible FP (pFP) that shares the characteristics of FPs of Class-I viral fusion proteins including high Ala/Gly content, intermediate hydrophobicity, few charged residues. To test the hypothesis that this region contains the CoV FP, we systemically mutated every residue in the pFP of Middle East Respiratory Syndrome betacoronavirus (MERS-CoV), and found that 11 of the 22 residues in the pFP (from G953 to L964, except for A956) were essential for S protein-mediated cell-cell fusion and virus entry. The synthetic MERS-CoV pFP core peptide ($_{955}$IAGVGWTAGL$_{964}$) induced extensive fusion of liposome membranes, while mutant peptide failed to induce any lipid mixing. We also selectively mutated residues in pFPs of two other β-CoVs, Severe Acute Respiratory Syndrome Coronavirus (SARS-CoV) and Mouse Hepatitis Virus (MHV). Although the amino acid sequences of these two pFPs differed significantly from that of MERS-CoV and each other, most of the pFP mutants of SARS-CoV and MHV also failed to mediate membrane fusion, suggesting that these pFPs are also the functional FPs. Thus, the FPs of 3 different lineages of β-CoVs are conserved in location within the S glycoproteins and in their functions, although their amino acid sequences have diverged significantly during CoV evolution.

Importance (150 words)

Within the Class-I viral fusion proteins of many enveloped viruses, the FP is the critical mediator of fusion of the viral envelope with host cell membranes leading to virus infection. FPs from within a virus family, like influenza viruses or human immunodeficiency viruses (HIV), tend to

share high amino acid sequence identity. In this study, we determined the location and amino acid sequences of the FPs of S glycoproteins of 3 β-CoVs: MERS-CoV, SARS-CoV, and MHV, and demonstrated that they were essential for mediating cell-cell fusion and virus entry. Interestingly, in marked contrast to the FPs of influenza and HIV, the primary amino acid sequences of the FPs of β-CoVs in 3 different lineages differed significantly. Thus, during evolution the FPs of β-CoVs have diverged significantly in their primary sequences, while maintaining the same essential biological functions. Our findings identify a potential new target for development of drugs against CoVs.

Introduction.

Viruses are obligate intracellular parasites, and host cell membranes act as a barrier to virus entry. Enveloped viruses initiate infection of cells through fusion of the viral and cellular membranes. CoVs are enveloped and single stranded plus sense RNA viruses that cause a variety of diseases among many different species (1). Phylogenetically, CoVs are divided into four genera: alphacoronavirus (α-CoV), betacoronavirus (β-CoV), gammacoronavirus (γ-CoV), and deltacoronavirus (δ-CoV) (2).

CoVs enter cells through the interactions of the viral S proteins with host receptors. Several cellular proteins have been identified as receptors for their respective CoVs. Specific examples include human angiotensin converting-enzyme 2 (hACE2) for SARS-CoV and human CoV NL63 (3, 4), human dipeptidyl peptidase IV (hDPP4) for MERS-CoV (5), bat DPP4 for bat CoV HKU4 (6), human aminopeptidase N (hAPN) for human CoV 229E (7), mouse carcinoembryonic antigen-related cell adhesion molecule 1a (mCEACAM1a) for MHV (8).

The CoV S protein is a Class-I viral fusion proteins. On the CoV virions, the 180-200 kDa S proteins are found as trimers. S monomers contain two subunits called S1 and S2. S1 contains the receptor binding domain (RBD) and is responsible for receptor recognition and binding, whereas S2 possesses the membrane fusion machinery (9, 10), including a fusion peptide (FP), two heptad repeat domains (called the N-terminal and C-terminal heptad repeats, HR-N and HR-C), the juxtamembrane domain (JMD) and a transmembrane domain (TMD) (Fig 1A).

To mediate membrane fusion, S protein must be activated, which requires both proteolytic cleavage (priming) and receptor binding with or without pH change (triggering) (11-13). Several host priming proteases are important for S protein mediated CoV entry, including

74 cathepsin B and L, serine protease TMPRSS2 and 4, trypsin, elastase, HAT, and furin (14-20). S
75 protein activation leads to a series of conformational changes and insertion of a putative FP into
76 target membrane, an essential step in membrane fusion and virus infection. Class-I viral fusion
77 proteins generally contain one FP, located either internally, like the FPs of the glycoprotein (Gp)
78 of Ebola virus and the envelope protein (Env) of avian sarcoma leukosis virus (ASLV) (21-24),
79 or immediately down stream of the "priming" site, as seen in the hemagglutinin (HA) of
80 influenza and the Env protein of HIV (25, 26). Although the primary sequences and lengths of
81 FPs vary significantly among different Class-I viral fusion proteins, they share several common
82 features. Most are rich in Ala and/or Gly, have an intermediate level of hydrophobicity with
83 membrane binding potential, form helical structures in the presence of trifluoroethanol (TFE),
84 and contain very few charged resides in the middle of their sequences (13, 25, 27).
85 Although significant efforts have been made to locate the FPs of different CoVs (28-35),
86 the exact locations and sequences of CoV FPs remains controversial. While Chambers et al
87 predicted that the CoV FP was likely adjacent to HR-N (ref), Manu et al proposed that the
88 sequence immediately following a critical and conserved trypsin cleavage site at the arginine of
89 position 797 (R797) of SARS-CoV S protein, SFIEDLLFNKVTLADAGF, may be the FP of
90 SARS-CoV S protein (32). In this study, we used bioinformatics to identify a 17-22 amino acids
91 long region, just upstream of HR-N, in S2 of different CoVs with characteristic features of the
92 FPs of other Class-I viral fusion proteins. Using mutational, biochemical, and biophysical
93 analyses of this region of the S proteins of 3 β-CoVs, MERS-CoV, SARS-CoV, and MHV, we
94 provide data to support this region as the functional FP of CoV S proteins.
95 **Materials and Methods**

96 **Cell culture.** HEK-293, 293T, HEK-293 cells stably expressing hACE2 (293/hACE2), HeLa
97 cells stably expressing hDPP4 (HeLa/hDPP4), and HeLa cells stably expressing mouse
98 CEACAM1a (HeLa/mCEACAM1a) were maintained in Dulbecco's modified Eagle's medium
99 (DMEM) (Invitrogen, Carlsbad, CA) supplemented with 10% fetal bovine serum (FBS) and 2%
100 penicillin-stereptomycin-fungizone (Invitrogen) at 37 °C with 5% CO_2.
101 **Constructs and mutagenesis.** The constructs, pcDNA-SARS-CoV SΔ19 (36), pcDNA-MERS-
102 CoV SΔ16 (37), and pcDNA-MHV S (38) have been described previously. Briefly, DNA
103 encoding codon-optimized SARS-CoV S protein lacking the last 19aa, or MERS-CoV S protein
104 lacking last 16aa but with a FLAG tag at the C-terminus, or full length MHV S protein was
105 cloned between *BamH I* and *Not I* sites of pcDNA3.1. All SARS-CoV, MERS-CoV, and MHV S
106 mutants were derived from the plasmid pcDNA-SARS-CoV SΔ19, pcDNA-MERS-CoV SΔ16,
107 and pcDNA-MHV S, respectively. All mutagenesis was carried out using Q5 mutagenesis kit
108 (NEB, MA, USA). After the entire coding sequences were verified by sequencing, the *BamH I*
109 and *Not I* containing mutated S gene was cloned back into pcDNA3.1. A plasmid encoding full-
110 length hACE2 (pACE2-cq) was kindly provided by M. Farzan (Scripps Research Institute,
111 Florida campus). A plasmids encoding full-length human DPP4 (pcDNA-hDPP4) was purchased
112 from Sino Biological Inc (Beijing, China). A plasmid encoding full-length mouse CEACAM1a
113 (mCEACAM1a) has been described previously (39). To express soluble human ACE2 (shACE2)
114 and soluble human DPP4 (shDPP4), DNA fragments encoding residues 19-615 of human
115 hACE2 with N-terminal 6his and FLAG tags and residues 40-766 of human DPP4 with C-
116 terminal 6his and AVI tags were cloned between *Sal I* and *Hind III* and between *BamH I* and
117 *Xho I* of modified pFASTBac1 vector with gp67 signal peptide, respectively. To express soluble
118 mouse CEACAM1a (smCEACAM1a), residues 1-236 of mCEACAM1a with C-terminal 6his

and AVI tags were cloned into *EcoR I* and *Not I* of pFASTBac1. These soluble receptors were expressed in High Five insect cells using the bac-to-bac system (Invitrogen) and purified using nickel affinity and ion-exchange chromatography.

Analysis of S protein expression on cell surface. Briefly, HEK-293T cells were transfected with 2 μg of either wild-type or mutant S protein-expressing plasmid using polyethyleneimine (PEI) (Polyscience Inc, Warrington, PA, USA). Forty hours later, cells were detached from plates by incubating with PBS+1mM EDTA for 5min at 37°C. After washing, cells were incubated with the respective primary anti-S antibody for 1 hour on ice. The primary antibodies for SARS-CoV SΔ19, MERS-CoV SΔ16, and MHV S protein were rabbit polyclonal anti-SARS S1 antibody (1:300 dilution) (Sinobiological Inc, Beijing, China), mouse monoclonal anti-MERS S antibody (1:300 dilution) (Sinobiological Inc, Beijing, China), and goat polyclonal anti-MHV S antibody (AO4) (1:200 dilution), respectively. After washing, cells were stained with Alexa Fluor 488 conjugated goat anti-rabbit IgG (1:200) (ZSGB-Bio LLC, Beijing, China) for SARS S, or goat anti-mouse IgG (1:200) (ZSGB-Bio LLC, Beijing, China) for MERS S, or rabbit anti-goat IgG (1:200) (ZSGB-Bio LLC, Beijing, China) for MHV S. After washing, cells were fixed with 1% paraformaldehyde and analyzed by flow cytometry.

Binding of soluble receptor. HEK-293T cells were transfected with plasmids encoding either wild-type or mutant S proteins with PEI. After 40 hours, cells were lifted with PBS+1mM EDTA and immediately washed twice with PBS+2% normal donkey serum (NDS). About $2x10^5$ cells were incubated with 1 μg of shACE2, or shDPP4, or smCEACAM1a for 1 hour on ice. After washing, cells were incubated with mouse monoclonal anti-FLAG M2 antibody (1:1,000 dilution) (Sigma, St Louis, MO, USA) for shACE2 and followed with Alexa Fluor 488 conjugated goat anti-mouse IgG (1:200), or rabbit polyclonal anti-AVI antibody (1:200 dilution) (Shanghai

Enzyme-linked Biotechnology Co., Shanghai, China) for shDPP4 and smCEACAM1a, and followed with Alexa Fluor 488 conjugated goat anti-rabbit IgG (1:200). Cells were fixed with 1% paraformaldehyde and analyzed by flow cytometry.

Production and transduction of S protein-pseudotyped lentiviruses. Pseudovirions with spike proteins were produced as described previously (40) with minor modifications. Briefly, plasmids encoding either wild-type or mutant S proteins were co-transfected into 293T cells with pLenti-Luc-GFP (a gift from Dr. Fang Li, Duke University) and psPAX2 (Addgene, Cambridge, MA) at a molar ratio of 1:1:1 by using PEI. The following day, the cells were fed with fresh medium. After 24 hrs incubation, the supernatant media containing pseudovirions were centrifuged at 800g for 5min to remove debris, and passed through a 0.45-μm filter. To quantify S protein-mediated entry of pseudovirions, susceptible cells were seeded at about 25-30% confluency in 24-well plates. The following day, cells were inoculated with 500ul of 1:1 diluted viruses. At 40 hours post-inoculation (PI), cells were lysed at room temperature with 120μl of media with an equal volume of Steady-glo (Promega, Madison, WI). Transduction efficiency was monitored by quantitation of luciferase activity using Modulus II Microplate Reader (Turner Biosystem, Sunnyvale, CA). All experiments were done in triplicate and repeated at least three times.

Detection of viral spike glycoproteins by western blot. To evaluate S protein expression in cells, HEK 293T cells were transfected with plasmids encoding either wild-type or mutant S proteins by using PEI. Forty hours later, cells were lysed with lysis buffer (50 mM Tris-HCl pH7.4, 150 mM NaCl, 2 mM EDTA, 1% Triton X-100, 0.1% SDS) containing protease inhibitors (Roche, USA). To determine S protein incorporation into pseudotype virions, the virion-containing supernatant was pelleted through a 20% sucrose cushion at 30,000 rpm at 4°C

for 2 h in a Beckman SW41 rotor (40). Viral pellets were resuspended into PBS. Cell lysates and pseudovirion pellets were separated on a 4-15% SDS-PAGE and transferred to a nitrocellulose blot. The SARS-CoV SΔ19, MERS-CoV SΔ16, and MHV S proteins were detected with polyclonal rabbit anti-SARS S1 antibodies (1:2,000), monoclonal mouse anti-MERS S antibody (1:1,000), and polyclonal goat anti-MHV S antibody (1:2,000), respectively, and the blots were further stained with horseradish peroxidase conjugated antibodies, respectively: goat anti-rabbit IgG (1:10,000), goat anti-mouse IgG (1:10,000), and rabbit anti-goat IgG (1:10,000), and visualized with Clarity Western ECL substrate (Bio-Rad, Hercules, CA, USA). The β-actin and HIV capsid protein (p24) were detected using mouse monoclonal anti-β-actin antibody (1:5,000) (Sigma, St Louis, MO, USA) and rabbit polyclonal anti-p24 antibody (1:5,000) (Sinobiological Inc, Beijing, China), respectively.

Cell-cell fusion assays. Cell-cell fusion assays were performed as previously described (37) with modifications. Briefly, 293T cells were co-transfected with plasmids encoding CoV S glycoprotein and GFP. Forty hours later, cells were detached with trypsin (0.25%) and overlaid on a 70% confluent monolayer of 293/hACE2, or HeLa/hDPP4, or HeLa/mCEACAM1a cells at a ratio of approximate one S-expressing cell to two receptor-expressing cells. After overnight incubation, images of syncytia were captured with a Nikon TE2000 epifluorescence microscope running MetaMorph software (Molecular Devices). To quantify S protein mediated cell-cell fusion, 293T cells were co-transfected with pFR-Luc, which contains a synthetic promoter with five tandem repeats of the yeast GAL4 binding sites that controls expression of the luciferase gene, and plasmid encoding S protein, and the receptor-expressing cells (293/hACE2, HeLa/hDPP4, or HeLa/mCEACAM1a) were transfected with pBD-NFκB, which encodes a fusion protein with DNA binding domain of GAL4 and transcription activation domain of NFκB.

188 The following day, S expressing 293T cells were lifted with trypsin and overlaid onto receptor
189 expressing cells at a ratio of about one S-expressing cell to two receptor-expressing cells. When
190 cell-cell fusion occurred, luciferase expression would be activated through binding of the GAL4-
191 NFκB fusion protein to GAL4 binding sites at the promoter of the luciferase gene. After 24 hrs
192 incubation, cells were lysed by adding 120µl of medium with an equal volume of Steady-glo, and
193 luciferase activity was measured with a Modulus II Microplate Reader. All experiments were
194 done in triplicate and repeated at least three times.
195 **Peptide synthesis.** All peptides were synthesized using a standard solid-phase FMOC (9-
196 fluorenylmethoxy carbonyl) method by Scilight Biotechnology LLC (Shanghai, China).
197 Purification was carried out by reversed-phase high-performance liquid chromatography (HPLC),
198 and verified by mass spectrometry. An Ahx-KKK linker was added to all peptides used in
199 circular dichroism (CD) spectroscopy analysis to increase peptide solubility in PBS. Peptides for
200 CD analysis include: CTRL: KWGQYTNSPFLTKGF-Ahx-KKK, a control peptide from a
201 previous SARS study (33); HIV FP (41): AVGIGALFLGFLGAAG-Ahx-KKK; and MERS pFP:
202 SSLLGSIAGVGWTAGLSSFAAI-Ahx-KKK. Peptides for lipid mixing study include: CTRL:
203 KWGQYTNSPFLTKGF; HIV FP: AVGIGALFLGFLGAAG; MERS short FP (sFP):
204 IAGVGWTAGL; MERS mutant FP (mFP): IAG**R**G**R**TAGL.
205 **CD spectroscopy.** CD spectroscopy analysis was performed to study the secondary structure of
206 fusion peptides in increasing trifluoroethanol (TFE) concentrations. CD spectra were acquired on
207 a Jasco J-815 spectropolarimeter (Jasco, Tokyo, Japan) using a 1-nm bandwidth with a 1-nm step
208 resolution from 195 to 260 nm at room temperature. Spectra were corrected by subtraction of its
209 respective solvent. The sample spectrum was smoothed with a Savitsky-Golay filter. The α-
210 helical content was estimated from the ellipticity value at 222nm, $[\theta]_{222,}$ according to the

empirical equation of Chen et al (42): %helical content=100*（$[\theta]_{222}$/-395000×(1-2.57/n)）, where *n* is the number of peptide bonds.

Preparation of liposomes. Equimolar amounts of egg phosphatidylethanolamine (PE), egg phosphatidylcholine (PC), and cholesterol (Avanti Polar Lipids, Alabaster, Ala., USA) were dried from chloroform into a thin film by constant flow of nitrogen gas, and rehydrated in Tris buffer (10 mM Tris, 150 mM NaCl, 0.1 mM EDTA, pH7.2) at a concentration of 10 mM. Large unilamellar vesicles (LUV) were prepared by the extrusion procedure (43). Briefly, after ten freeze-thaw cycles, liposomes were extruded 21 times through two stacked polycarbonate membranes with a pore size of 0.1 μm using an Avanti mini-extruder. Liposome with 0.6% (molar ratio) fluorescent resonance energy transfer (FRET) pairs Rho-PE and NBD-PE (Thermo Fisher) were prepared in the same way.

Lipid mixing. Lipid mixing was determined using the resonance energy transfer assay, described by Struck et al (44) with minor modifications. Briefly, Rho-PE and NBD-PE labeled liposomes were mixed with unlabeled liposomes at a ratio of 1:9. The final lipid concentration was 300 μM. Specified amounts of various peptides were added to initiate fusion, and changes in fluorescence were monitored at 535 nm with the excitation wavelength set at 465 nm and a slit width of 4 nm using Fluromax-4 (Horiba, Paris, France). The initial residual fluorescence of the labeled and unlabeled vesicles was set up as baseline for 0% fluorescence value (f_0); 100% fluorescence value (f_{100}) was achieved by addition of Triton X-100 to final concentration of 0.2%. The extent of lipid mixing was calculated using the following formula: $\%F_t=(f_t-f_0)/(f_{100}-f_0)*100$, where f_t is the fluorescence value observed after addition of fusion peptide at time t.

Results

Journal of Virology

233 During membrane fusion, the FP of S proteins inserts into the host membranes. We
234 reasoned that CoV FPs might share some common properties with the transmembrane domains
235 (TMD) and that the location of the FP within the S protein might be predictable by using TMD
236 prediction software programs. The FPs of HIV-1 Env and influenza HA have been studied
237 extensively and their locations and amino acid sequences are known. As a proof of concept, we
238 first tested whether TM software programs could accurately identify the FPs of HIV-1Env and
239 influenza H1N1HA proteins. Both the FPs and TMDs of HIV-1 Env and influenza HA were
240 accurately identified by two software programs, TMpred
241 (http://www.ch.embnet.org/software/TMPRED_form.html) and TMHMM
242 (http://www.cbs.dtu.dk/services/TMHMM/) (Data not shown). Subsequently, we applied these
243 two software programs to analyze S proteins of a wide variety of CoVs. The positions of the
244 TMDs of the S proteins of all CoVs studied were correctly identified by both software programs
245 (Fig 1B). In addition, both of these TMD prediction programs identified another region
246 consistently flanked by YT at the N-terminus and PF at the C-terminus in all of the S proteins of
247 the CoVs tested (Figs. 1B and 1C). Although the primary amino acid sequences of this region
248 were not conserved in all of the CoVs studied, they were all Ala or Gly rich, relatively
249 hydrophobic, and contained no charged residues, characteristics shared by the FPs of other
250 Class-I viral fusion proteins (Fig. 1C). We named this region in CoV S proteins the possible FP
251 (pFP).

252 To investigate if the pFP is the functional fusion peptide of CoVs, we selected the S
253 protein of MERS-CoV, a lineage C β-CoV, as an example. The MERS-CoV pFP contains amino
254 acids 949 to 970 (Fig. 1C). Individual and occasionally double amino acid substitutions were
255 introduced at each position of pFP (Fig. 1D). First, we determined if any of the mutations altered

256 the expression of S protein in 293T cells. Consistent to our previous report (37), two bands
257 around 200 kDa were detected in the cell lysate of 293T cells expressing wild-type (WT) S
258 protein, likely reflecting the different glycosylation of full length S proteins during transport
259 through the Golgi apparatus. However, the cell lysate also contained a significant proportion of S
260 protein cleaved between S1 and S2, around 100 kDa, which was absent in our previous report,
261 but previously reported by the Pohlmann laboratory (45). The difference between this study and
262 our early report likely resulted from different culture conditions, especially sera and media from
263 different vendors. Among the total 44 G, A, V, or R substitutions, 30 (S949G, S950G, L952A,
264 G953A, G953R, S954G, S954R, I955G, A956V, A956R, G957A, G957R, V958G, V958R,
265 I955G/V958G, G959A, G959R, W960G, W960R, V958G/W960G, T961A, A962V, A962R,
266 G963A, G963R, L964G, L964R, S965G, S966G, and A968V) showed no or minor effects on S
267 protein expression or processing when compared to WT(Fig. 2A and Table 1). On the contrast,
268 14 substitutions (L951G, L952G, L951G/L952G, S965R, S966R, F967G, L964F/F967G, A968R,
269 A969V, A969R, I970G, P971V, F972G, and I970G/F972G) showed significant reductions in S
270 protein expression and changes in patterns of S protein processing (Fig. 2A and Table 1). The
271 cleaved S protein species were almost absent in corresponding cell lysates, suggesting that these
272 residues (L951, L952, S965, S966, F967, A968, A969, I970, P971, and F972) may be important
273 for S protein folding and processing.

274 We then investigated if any amino acid substitutions in the pFP influenced transport of
275 the S protein to the cell surface. The 293T cells expressing WT or mutant S proteins were
276 incubated on ice with mouse monoclonal anti-MERS-CoV S protein antibody and analyzed by
277 flow cytometry. The same 30 mutants that showed WT levels of S protein expression in cell
278 lysates also showed WT levels of S protein on the cell surface (Fig. 2B and Table 1). As

279 expected, the mutants with defects in S protein expression and processing also showed only low
280 levels of S proteins on the cell surface.

281 Although the pFP is located within the MERS-CoV S2 subunit, amino acid substitutions
282 in pFP might affect S protein binding to its cognate receptor, hDPP4, by altering the overall
283 conformation of the S protein. To determine whether or not any amino acid substitution in pFP
284 changed S protein binding to hDPP4, we used V5-tagged soluble hDPP4 (shDPP4) to bind 293T
285 cells transiently expressing WT or pFP mutant S proteins of MERS-CoV. The percentage of cells
286 that bound shDPP4 and the level of shDPP4 bound to S protein were quantitated by flow
287 cytometry. The same 30 mutant S proteins that showed WT levels of expression on cell surface
288 also bound to shDPP4 at levels similar to WT S protein (Fig. 3 and Table 1), indicating that these
289 pFP mutations had no effect on receptor binding.

290 Because the fusion peptide is essential for S protein-mediated membrane fusion, we then
291 explored whether any mutation in pFP altered MERS-CoV S protein-mediated cell-cell fusion.
292 To more easily visualize cell-cell fusion or syncytia, the 293T cells expressing S protein were co-
293 transfected with a GFP-expressing plasmid, then overlaid with HeLa/hDPP4 cells in the presence
294 of trypsin. Consistent with our previous report (37), WT MERS-CoV S protein induced very
295 large syncytia (Fig 4) and syncytia formation depended on the availability of hDPP4 (data not
296 shown). Among 30 pFP S protein mutants that were expressed well, transported to the cell
297 surface efficiently, and bound to hDPP4 at levels similar to WT, 14 mutants (S949G, S950G,
298 G953A, S954G, A956V, A956R, G957A, V958G, G959A, A962V, G963A, L964G, S965G, and
299 A968V) induced large syncytia in HeLa/hDPP4 cells similar to WT, while 12 mutants (G953R,
300 S954R, I955G, G957R, V958R, I955G/V958G, W960R, V958G/W960G, T961G, A962R,
301 G963R, and L964R) induced little or no syncytia formation, and 4 mutants (L952A, G959R,

W960G, and S966G) induced syncytia of much smaller size than WT (Fig 4). These results indicate that these 13 residues, L952, G953, S954, I955, G957, V958, G959, W960, T961, A962, G963, L964, and S966, in MERS-CoV S protein are critical for S protein-mediated, receptor-dependent membrane fusion that would lead to virus infection.

To quantify the effect of amino acid substitutions on S protein-mediated syncytia formation, we utilized a luciferase-based quantification assay from a yeast two hybrid system from Stratagene-Agilent Technologies, Inc. Compared to mock transfection and parental HeLa cell controls, fusion of 293T cells expressing WT MERS-CoV S proteins with HeLa/hDPP4 cells increased luciferase activity by about 1,000-fold (Fig.5). The overall pattern of cell-cell fusion induced by pFP mutants in this quantification assay was very similar to our visual method (Fig. 4 and 5, Table 1). Among the same 30 mutants showing WT level of expression and receptor binding, 16 mutants (S949G, S950G, L952A, G953A, S954G, A956V, A956R, G957A, V958G, G959A, A962V, G963A, L964G, S965G, S966G, and A968V) retained 50-110% of WT level fusion activity, but 14 mutants (G953R, S964R, I955G, G957R, V958R, I955G/V958G, G959R, W960G, W960R, V958G/W960G, T961G, A962R, G963R, and L964R) reduced S protein-mediated cell-cell fusion by more than 85%, indicating that these residues (G953, S954, I955, G957, V958, G959, W960, T961, A962, G963, and L964) are essential for membrane fusion.

To determine whether or not any mutation in the pFP of the S protein of MERS-CoV also affected virus entry, we measured transduction of HeLa/hDPP4 cells by lentiviral pseudovirions with envelopes containing either WT or pFP mutant MERS-CoV S proteins. Compared to mock control (pseudovirions without any S protein), the luciferase activity in HeLa/hDPP4 cells increased by more than 10,000 fold following transduction by pseudovirions with WT MERS-CoV S proteins (Fig. 6A). Among the same 30 mutants that showed little or no effects on S

protein expression or receptor binding (Figs 2A, 2B, 3, and Table 1), 5 mutants (L952A, G953A, G953R, G963R, and S966G) showed marked reduction in S protein incorporation into pseudovirions, whereas the S proteins of the other 25 mutants were incorporated into pseudovirions as well as WT S protein (Fig. 6B). Ten out of these 25 amino acid substitutions, S954R, I955G, G957R, V958R, I955G/V958G, W960R, V958G/W960G, T961G, A962R, or L964R, almost abolished MERS-CoV S protein-mediated, receptor-dependent pseudovirion entry (Fig 6A and Table 1), suggesting that S954, I955, G957, V958, W960, T961, A962, and L964 are essential for virus entry. In addition, G959R mutation also reduced the transduction by more than 95%, indicating that G959 may also be critical for virus entry too (Fig. 6A). Interestingly, although G953A, G953R, and G963R mutants showed reduced but similar levels of S protein incorporation into pseudovirions (Fig. 6B), the infectivity of the pseudovirions differed drastically. While G953A result in only 30% of WT level of pseudovirion entry, the G953R and G963R mutations almost abrogated S protein mediated pseudovirion entry, indicating that G953 and G963 may also be important for virus entry.

Because the FPs of most Class-I viral fusion proteins fold predominantly in an α-helix structure in the presence of TFE (13), we used circular dichroism spectroscopy (CD) analysis to investigate whether our MERS-CoV pFP also adapts an α-helical fold. A scrambled peptide from a previous SARS-CoV study (33) was chosen as the negative control, and the FP of HIV-1 was selected as the positive control (46). To facilitate the synthesis of the peptides and increase their solubility, an aminocaproic acid (Ahx) linker followed by 3 Lys residues (Ahx-KKK) was added to the C-termini of the peptides. Consistent with the previous reports (46), while the FP of HIV-1 folded as a random coil in Tris/salt buffer, it formed an α-helix in the presence of trifluoroethanol (TFE) (Fig. 7A), a solvent known to stabilize the α-helical formation (47).

Similarly, in the absence of TFE, the pFP of MERS-CoV (SSLLGSIAGVGWTAGLSSFAAI) folded as a random coil, but with the addition of TFE, it folded as an α-helix. At 95% of TFE, helical population accounted for more than 64% (Fig. 7A).

FPs of Class-I viral fusion proteins also promote membrane fusion when mixed with liposomes. Accordingly, we investigated whether the pFP of MERS-CoV S protein could mediate liposome fusion using a FRET-based assay. To rule out any possible effect of the Ahx-KKK tag, we decided to use peptides without any tag. However, because of the technical difficulty of synthesizing the full length pFP without the AHX-KKK tag, we decided to use instead the core sequence of pFP ($_{955}$IAGVGWTAGL$_{964}$, called "short pFP" or sFP) in this study, in which almost all of the residues were shown to be essential for cell-cell fusion and virus entry. As shown in Fig 7B, both the FP of HIV-1 and the sFP of MERS-CoV induced membrane fusion of liposome in a concentration dependent manner, whereas the negative control peptide did not induce any significant lipid mixing. Moreover, when we replaced V958 and W960, two residues essential for cell-cell fusion and virus entry, with Arg in the MERS-CoV sFP peptide, the resulting mutant FP (mFP) ($_{955}$IAG**R**G**R**TAGL$_{964}$) failed to induce any noticeable lipid mixing, confirming that these two residues are essential for lipid mixing.

Having established the essential roles in membrane fusion and virus entry of the pFP of the S protein MERS-CoV, a β-CoV in group C, we also investigated the functional role of the pFPs of other CoVs. After examining the alignment of the pFPs of different CoVs (Fig. 1B), we selected the pFPs of the S proteins of SARS-CoV, a lineage B β-CoV, and MHV, a lineage A β-CoV, for functional study. While the pFP of SARS-CoV shares the same length and has about 1/3 of amino acid sequence identity with the pFP of MERS-CoV, the pFP of MHV differs markedly from that of MERS-CoV in both length and amino acid sequence. Since hydrophobic

residues in the pFP of MERS-CoV play important roles in membrane fusion, we selected W868, F870, L876 and I878 of SARS-CoV S protein and M936, F937, P938, P939, and W940 of MHV S protein for further analysis. Single Arg and/or Gly substitutions were introduced into the MHV and SARS-CoV S proteins at these positions.

With the exception of I878-related mutants, the pFP mutant S proteins of SARS-CoV were expressed well (data not shown), bound well to its receptor, hACE2, at levels similar to WT (data not shown), and were incorporated into pseudovirions efficiently (Fig 8B). I878 mutants (I878G, I878R, and double mutant L876G/I878G) were expressed slightly less well in cell lysates (data not shown) and showed reduced S protein incorporation into pseudovirons (Fig. 8B), indicating that I878 may play a role in folding and transport of S protein. Similar to MERS-CoV S protein, all Arg mutations in pFP of SARS-CoV effectively abolished S protein mediated cell-cell fusion and virus entry (Fig. 8A, 8C, and Table 1), suggesting that these residues are indeed essential for membrane fusion. Compared to Arg mutations, Gly substitutions in the pFP of SARS S protein had less effect on cell-cell fusion and virus entry. Interestingly, although the single mutants, W868G and F870G, showed almost WT level infection, the double mutant W868G/F870G abolished S protein mediated virus entry (Fig. 8A), confirming that these two residues in S protein of SARS-CoV are important for membrane fusion.

All MHV S protein pFP with single Arg substitutions (M936R, F937R, P938R, P939R, and W940R) showed significant reduction in both S-mediated pseudovirion entry (Fig. 8D and Table 1) and cell-cell fusion (Fig. 8F and Table 1). S proteins with M936R substitutions, however, showed significantly decreased expression of S protein in cell lysate (data not shown) and incorporation into pseudovirions (Fig. 8E). This may partly explain why M936R mutations had detrimental effects on virus infection and cell-cell fusion. P938R substitution also showed

394 slight reduction in expression and virion incorporation of S protein. In contrast, S proteins with
395 F937R, P939R, and W940R substitutions had wild-type levels of S protein expression (data not
396 shown) and incorporation into virions (Fig. 8E), and binding to its cognate receptor (data not
397 shown), mCEACAM1a, but failed to mediate virus entry or syncytia formation. These data
398 indicate that F937, P939, and W940 in the pFP may be essential for MHV S protein-mediated
399 membrane fusion.

400 **Discussion.**

401 Proteolytic priming is one of the early essential steps required to activate the fusion
402 potential of Class-I viral fusion proteins, and is believed to release the restrain on the viral FP
403 leading to exposure of the FP. The proteolytic priming sites for most of the Class-I viral fusion
404 proteins are either immediately proximal to or not far upstream of the viral FP (21-26). Therefore,
405 identifying the key proteolytic priming site may lead to discovery of a viral FP. However, in the
406 case of CoVs, the priming sites are less clear. In an attempt to identify the trypsin cleavage site
407 essential for MERS-CoV S protein mediated trypsin-dependent entry, we mutated several trypsin
408 sites (R884G/R887G, K897G, R921G, and K933G) upstream of the N-terminus of HR-N of
409 MERS-CoV S protein (48, 49). Surprisingly, we found that none of these sites was essential for
410 trypsin-primed MERS-CoV S protein-mediated virus entry (Data not shown). Therefore, there
411 might be built-in redundancy of trypsin priming sites within the MERS-CoV S protein such that
412 cleavage by trypsin might occur at multiple sites and single cleavage at any one of these sites
413 might be sufficient to prime the MERS-CoV S protein.

414 Since there was not a single essential trypsin priming site for the S protein of MERS-CoV,
415 we used an alternative approach to look for the FP of MERS-CoV S protein. Using TMpred and
416 TMHMM software programs to analyze the S2 domains of a variety of CoVs, we identified a

417 region in S2 that is flanked by YT at the N-terminus and PF at the C-terminus and found in all
418 the CoVs studied (Figs 1B and 1C). This pFP region has characteristics of the known FPs of
419 other Class-I viral fusion proteins, Gly or Ala rich, relatively hydrophobic, and without charged
420 residues. This pFP region is located at about 7-23 amino acids upstream of the N-terminus of
421 HR-N of CoV S proteins, depending on where the N-terminus of HR-N was proposed (48-53).
422 Mutagenesis analysis on the pFPs of MERS-CoV, SARS-CoV, and MHV S proteins revealed
423 that this region was essential for S protein mediated syncytia formation and virus entry (Table 1),
424 and strongly support the idea that the pFP of β-CoV S protein is the functional viral fusion
425 peptide. This conclusion is further strengthened by our findings that the synthetic pFP of MERS-
426 CoV S protein formed an α-helix in the presence of TFE and its core short sequence, called sFP,
427 mediated membrane fusion of liposome efficiently (Fig 7), which are characteristics of FPs of
428 other Class-I viral fusion proteins (13). Our results are also consistent with previous biophysical
429 studies on synthetic peptides from SARS-CoV S protein (29, 33) and previous studies in MHV
430 showing that P939 may be critical for membrane fusion and virus infection (54, 55).
431 About one third of the residues located at the C-terminus of the pFP of MERS-CoV S
432 protein appear to play important roles in the stability and processing of the S protein, since
433 introduction of amino acid substitutions into these positions significantly reduced S protein
434 expression, processing and incorporation into pseudotyped virions. Residues close to the C-
435 terminus of the pFP of the SARS-CoV S protein also appear to be important for S protein folding,
436 as replacement of I878 with R or G also decreased S protein expression and incorporation into
437 virions. However, this region might also be important for membrane fusion mediated by S
438 protein. A recent study on SARS-CoV by Liao et al (56) raised the possibility that this region
439 might make direct interactions with the JMD in S protein during membrane fusion.

Among the amino acid substitutions that we introduced into the pFP of MERS-CoV S protein, Arg had a more profound effect on the function of the pFP of MERS-CoV S protein than Gly, Ala, or Val. Compared to Gly, Ala, and Val, Arg is positively charged and its side chain is significant longer than Gly, Ala, or Val, hence Arg substitution represents a more dramatic change than these amino acids. Moreover, Arg substitution may cause a higher free energy barrier for insertion of FP into membrane (57). Although the exact mechanism(s) of how these substitutions in the pFP abrogate membrane fusion requires further investigation, there are several possibilities. Introduction of mutation(s) into the pFP of MERS-CoV S protein might distort the structure of FP required for membrane fusion similar to G1V and W14A mutations of the FP of influenza HA (58-60). Alternatively, the substitutions might change how the FP inserts into membranes (61-63), or affect the oligomerization of the FPs that is important for membrane fusion (64, 65).

Recent studies in influenza HA (66), paramyxovirus F protein (67), and HIV Env (68) reveal that many viral FPs interact and oligomerize with their TMDs in the lipid, which promotes lipid mixing and membrane fusion. Whether the FP and TMD of CoV S protein interact with each other during membrane fusion remains to be further determined. Interestingly, the primary amino acid sequences of the TMDs among different CoVs also do not share high identity (Fig 9). Of note, there is a GXXXG or (small)XXX(small) motif (G, Gly; small, Ala or Gly or Ser; X, any residue) present in all of the pFPs of CoVs. These motifs were initially discovered in human glycophorin A and have subsequently been implicated in TMD interactions of more than 20 proteins (69). Recent studies in influenza HA and HIV Env have suggested that such GXXXG motifs may also play an important role in FP:FP or FP:TMD interaction (66, 68, 70). There are two GXXXG motifs, GSIAG and GWTAG, within the FP of MERS-CoV. Replacement in

463 MERS-CoV S protein of any one of these four Gly residues (G953, G957, G959, or G963) with
464 Arg abrogated the membrane fusion activity of the viral protein. However, whether these
465 GXXXG motifs in the pPF of MERS S protein are essential for oligomerization or interaction
466 with the TMD requires further investigation.

467 FPs of some Class-I viral fusion proteins, like HIV Env and influenza HA, share high
468 identity in primary amino acid sequence within each virus family. In marked contrast, this study
469 found no strong amino acid sequence identity among the pFPs of MERS-CoV, SARS-CoV, and
470 MHV. The lengths of the FPs of these three different lineages of β-CoVs also differ significantly,
471 ranging from 18 for MHV to 22 amino acids for MERS-CoV and SARS-CoV (Fig. 1B). Within
472 each lineage of β-CoVs, the pFPs appear to be better conserved (Fig 1B). Although underline
473 mechanism(s) causing the amino acid sequence diverge of FPs of different lineages of β-CoVs
474 remains to be determined, CoV RNA-dependent RNA polymerase error, recombination, and
475 selective pressure during evolution likely contribute to these changes. Previous study of MHV
476 persistent infection in DBT cells showed that accumulation of mutations in fusion peptide and
477 HR-N could lead to extending host range (55). The lack of conservation of the pFP amino acid
478 sequences, however, is not unique for CoVs, as FPs from different paramyxoviruses also lack
479 high identity in their primary amino acid sequences (67).

480 As an internal fusion peptide, how does the activated FP of CoVs fold and mediate
481 membrane fusion? Recent studies have demonstrated that FPs from different Class-I viral fusion
482 proteins might adapt different conformations to mediate membrane fusion. Depending on the
483 lipid composition, the FPs of HIV-1 Envs and PIV F proteins can fold as either α–helix (67, 71)
484 or β–sheet (65, 72), and both can be fusiogenic. In contrast, the overall conformation of the FPs
485 of Ebola Gp and influenza HA is α–helical in the presence of TFE, but they fold as hairpin-like

structure or "knuckle" conformations when they insert into their target membranes (63, 73). Sequence analysis of the S proteins of different CoVs (Fig 1B) shows the presence in the pFPs of a Gly-Gly (GG) motif in α-, γ-, and δ-CoVs or a Pro-Pro (PP) motif in β-CoVs in lineage A. As GG and PP motifs favor the formation of turn or hairpin structures, this observation suggests that the FPs of some CoVs might also adapt a hairpin-like structure when inserting into host membranes. In the FPs of SARS-CoV and MERS-CoV in β-CoV groups b and c respectively, however, neither a GG nor a PP motif is present. Of note, the FP from group 2 influenza HA also lacks a central GG or PP motif, but instead it forms a hairpin-like structure with G13 at the turn with a Trp and a hydrophilic residue immediately following G13 (74). Interestingly, a similar motif is also present in the pFPs of SARS-CoV and MERS-CoV (Fig 1B).

While all known Class-III viral fusion protein have two fusion loops, all known Class-I viral fusion proteins except for CoV S protein only have a single fusion peptide. In the case of CoVs, in addition to pFP found in this study, Manu et al previously found a highly conserved region in SARS-CoV S protein essential for membrane fusion and proposed it as the possible fusion peptide (32), although this sequence lacks some common features of FPs of other Class-I viral fusion proteins, including high Ala/Gly content. Their proposed FP is about 80 amino acids away from the N-terminus of HR-N (50, 52) and about 40 amino acids upstream of the N-terminus of our pFP. The possibility of presence of two possible fusion peptides in the S protein of CoV is very intriguing. How these two possible fusion peptides collaborate to mediate membrane fusion requires further investigation.

In summary, using a bioinformatics approach we have identified a region in the S proteins of CoVs that has several properties the FPs of several classical Class-I viral fusion proteins. Further molecular biological, biochemical, and biophysical analyses demonstrated that

this region is essential for receptor-dependent membrane fusion mediated by S proteins of several β-CoVs in different lineages, strongly suggesting that it is the functional FP of these and likely other CoVs. These findings will provide significant clues for future studies of the membrane fusion mechanism of CoVs and may provide a new target for drugs against CoV infections.

Acknowledgement

This work was supported by grants from Chinese Science and Technology Key Projects (2014ZX10004001), National Natural Science Foundation of China (31470266), MOHRSS of China (9019005), and Institute of Pathogen Biology, CAMS (2014IPB101 and 2015IPB301) to ZQ. This work was also supported by PUMC Youth Fund and the Fundamental Research Funds for the Central Universities (3332013118), and the Program for Changjiang Scholars and Innovative Research Team in University (IRT13007).

Figure legend.

Figure 1. pFPs of CoVs. (A) Diagram of CoV spike protein. NTD, N-terminal domain; C-domain, C-terminal domain; Cleavage site, protease cleavage site separating S1 and S2; pFP, possible fusion peptide; HR-N, N-terminal heptad repeat; HR-C, C-terminal heptad repeat; JMD, Juxtamembrane domain; TMD, transmembrane domain. (B) Locations of pFPs and TMDs of S proteins of representative CoVs predicted by TMPred. (C) Amino acid sequence alignment of the pFPs of different CoVs. (D) Summary of the amino acid substitutions made in the pFP of MERS-CoV S protein.

Figure 2. Analysis of expression of pFP mutants of MERS-CoV S protein in 293T cells. (A) Western blot analysis of expression of WT or mutant MERS S protein in cell lysate. The MERS S protein was detected by using mouse monoclonal anti-MERS S antibody; β-actin was detected

with mouse monoclonal anti-actin antibody. (B) Analysis of surface expression of mutant MERS-CoV S protein by flow cytometry. MERS-CoV S protein expressing 293T cells were stained with mouse monoclonal anti-MERS S antibody. The amount of wild-type S protein on cell surface was set as 100%. All of the experiments shown were repeated at least three times.

Figure 3. Receptor binding by mutant MERS S proteins. MERS-CoV S protein expressing 293T cells were incubated with soluble AVI-tagged hDPP4, followed with polyclonal rabbit anti-AVI antibody and FITC conjugated goat anti-rabbit IgG. The results from wild-type were set as 100%.

Figure 4. Cell-cell fusion mediated by WT or mutant MERS-CoV S protein. MERS-CoV S protein expressing 293T cells were transiently transfected with eGFP, then incubated with HeLa/hDPP4 cells for overnight in the presence of trypsin.

Figure 5. Quantitative analysis of syncytia formation mediated by WT or mutant MERS-CoV S protein. Cell-cell fusion was quantified by measurement of luciferase activities. Typically, the relative luciferase activities from cell-cell fusion induced by wild-type S protein were over 10^7, while the reading for mock control was less than 1000. The experiments were done at least three times.

Figure 6. Entry of pseudotype virions with wild-type or mutant MERS S protein. A. Entry of pseudovirions with wild-type or mutant MERS-CoV S proteins into HeLa/hDPP4 cells. Pseudovirus entry was quantitated by luciferase activity at 40 hrs post inoculation. A typical transduction by wild-type S protein pseudoviruses resulted in increase of over 10,000-fold of luciferase activity. The experiments were repeated at least three times and an average of three experiments is shown. B. Detection of wild-type or mutant S protein incorporation into pseudovirions by western blot analysis. MERS S protein was detected using mouse monoclonal anti-MERS S antibody; p24, a gag protein of HIV, was detected using rabbit polyclonal anti-p24

antibodies. FL S: full length S protein. The experiments were repeated twice and a representative is shown.

Figure 7. Biophysical analysis of synthetic pFP peptide of MERS-CoV. A. CD analysis of secondary structure of pFP of MERS-CoV S protein. CTRL: KWGQYTNSPFLTKGF-Ahx-KKK, a control peptide from previous SARS-CoV peptide study (33); HIV FP: AVGIGALFLGFLGAAG-Ahx-KKK; MERS pFP: SSLLGSIAGVGWTAGLSSFAAI-Ahx-KKK. All peptides were dissolved in PBS, and their CD spectrum was measured in the presence of indicated concentration of TFE. Experiments were done twice and one representative is shown. B. Lipid mixing induced by synthetic pFP of MERS-CoV S protein. LUVs were made with equal moles of PE, PC, and cholesterol. The extent of lipid mixing was determined by monitoring the changes in fluorescence intensity at 535 nm at 37°C upon addition of peptide. Each data point is averaged from three independent experiments, and error bars represent standard deviations of the means. CTRL: KWGQYTNSPFLTKGF; HIV FP: AVGIGALFLGFLGAAG; MERS sFP: IAGVGWTAGL; MERS mFP: IAGRGRTAGL.

Figure 8. Effects of mutations at the pFPs of SARS-CoV and MHV on pseudovirus transduction and cell-cell fusion. A, D. Entry of wild-type or mutant SARS-CoV S protein pseudovirions into 293/hACE2 cells (A) or MHV S protein pseudovirions into HeLa/mCEACAM1a cells (D). Pseudovirus entry was quantitated by luciferase activity at 40 hrs post inoculation. The experiments were repeated at least three times and average of three experiments is shown. B, E. Detection of wild-type or mutant S protein of SARS-CoV (B) or MHV (E) incorporation into pseudovirions by western blot analysis. SARS S protein was detected using rabbit polyclonal anti-SARS S1 antibody; MHV S protein was detected using goat polyclonal anti-MHV S antibody AO4; p24, a gag protein of HIV, was detected using rabbit polyclonal anti-p24

antibodies. The experiments were repeated at least three times and one representative is shown. C, F. Cell-cell fusion mediated by mutant SARS (C) or MHV (F) S proteins. Experiments were performed as in Fig. 3B, except that 293/hACE2 cells were used as targets for SARS-CoV S protein (C) and HeLa/mCEACAM1a cells were used as targets for MHV S protein (F). An average of three experiments is shown.

Figure 9. Alignment of TMDs of S proteins of representative CoVs.

Journal of Virology

585 Reference

586 1. **Masters PS, Perlman S.** 2013. Coronaviridae, p. 825-858. *In* Knipe DM, Howley PM
587 (ed.), Fields Virology, sixth ed, vol. 1.
588 2. **Viruses ICoTo.** 2011. Virus Taxonomy: 2011 Release.
589 http://ictvonline.org/virusTaxonomy.asp?version=2011.
590 3. **Hofmann H, Pyrc K, van der Hoek L, Geier M, Berkhout B, Pohlmann S.** 2005.
591 Human coronavirus NL63 employs the severe acute respiratory syndrome coronavirus
592 receptor for cellular entry. Proc Natl Acad Sci U S A **102:**7988-7993.
593 4. **Li W, Moore MJ, Vasilieva N, Sui J, Wong SK, Berne MA, Somasundaran M,**
594 **Sullivan JL, Luzuriaga K, Greenough TC, Choe H, Farzan M.** 2003. Angiotensin-
595 converting enzyme 2 is a functional receptor for the SARS coronavirus. Nature **426:**450-
596 454.
597 5. **Raj VS, Mou H, Smits SL, Dekkers DH, Muller MA, Dijkman R, Muth D, Demmers**
598 **JA, Zaki A, Fouchier RA, Thiel V, Drosten C, Rottier PJ, Osterhaus AD, Bosch BJ,**
599 **Haagmans BL.** 2013. Dipeptidyl peptidase 4 is a functional receptor for the emerging
600 human coronavirus-EMC. Nature **495:**251-254.
601 6. **Yang Y, Du L, Liu C, Wang L, Ma C, Tang J, Baric RS, Jiang S, Li F.** 2014.
602 Receptor usage and cell entry of bat coronavirus HKU4 provide insight into bat-to-human
603 transmission of MERS coronavirus. Proc Natl Acad Sci U S A **111:**12516-12521.
604 7. **Yeager CL, Ashmun RA, Williams RK, Cardellichio CB, Shapiro LH, Look AT,**
605 **Holmes KV.** 1992. Human aminopeptidase N is a receptor for human coronavirus 229E.
606 Nature **357:**420-422.
607 8. **Williams RK, Jiang GS, Holmes KV.** 1991. Receptor for mouse hepatitis virus is a
608 member of the carcinoembryonic antigen family of glycoproteins. Proc Natl Acad Sci U
609 S A **88:**5533-5536.
610 9. **Babcock GJ, Esshaki DJ, Thomas WD, Jr., Ambrosino DM.** 2004. Amino acids 270
611 to 510 of the severe acute respiratory syndrome coronavirus spike protein are required for
612 interaction with receptor. J Virol **78:**4552-4560.
613 10. **Wong SK, Li W, Moore MJ, Choe H, Farzan M.** 2004. A 193-amino acid fragment of
614 the SARS coronavirus S protein efficiently binds angiotensin-converting enzyme 2. J
615 Biol Chem **279:**3197-3201.
616 11. **Li F, Berardi M, Li W, Farzan M, Dormitzer PR, Harrison SC.** 2006.
617 Conformational states of the severe acute respiratory syndrome coronavirus spike protein
618 ectodomain. J Virol **80:**6794-6800.
619 12. **Harrison SC.** 2015. Viral membrane fusion. Virology **479-480:**498-507.
620 13. **White JM, Delos SE, Brecher M, Schornberg K.** 2008. Structures and mechanisms of
621 viral membrane fusion proteins: multiple variations on a common theme. Crit Rev
622 Biochem Mol Biol **43:**189-219.
623 14. **Bertram S, Glowacka I, Muller MA, Lavender H, Gnirss K, Nehlmeier I, Niemeyer**
624 **D, He Y, Simmons G, Drosten C, Soilleux EJ, Jahn O, Steffen I, Pohlmann S.** 2011.
625 Cleavage and activation of the severe acute respiratory syndrome coronavirus spike
626 protein by human airway trypsin-like protease. J Virol **85:**13363-13372.
627 15. **Millet JK, Whittaker GR.** 2014. Host cell entry of Middle East respiratory syndrome
628 coronavirus after two-step, furin-mediated activation of the spike protein. Proc Natl Acad
629 Sci U S A **111:**15214-15219.

16. **Belouzard S, Madu I, Whittaker GR.** 2010. Elastase-mediated activation of the severe acute respiratory syndrome coronavirus spike protein at discrete sites within the S2 domain. J Biol Chem **285:**22758-22763.
17. **Simmons G, Reeves JD, Rennekamp AJ, Amberg SM, Piefer AJ, Bates P.** 2004. Characterization of severe acute respiratory syndrome-associated coronavirus (SARS-CoV) spike glycoprotein-mediated viral entry. Proc Natl Acad Sci U S A **101:**4240-4245.
18. **Watanabe R, Matsuyama S, Shirato K, Maejima M, Fukushi S, Morikawa S, Taguchi F.** 2008. Entry from the cell surface of severe acute respiratory syndrome coronavirus with cleaved S protein as revealed by pseudotype virus bearing cleaved S protein. J Virol **82:**11985-11991.
19. **Bertram S, Dijkman R, Habjan M, Heurich A, Gierer S, Glowacka I, Welsch K, Winkler M, Schneider H, Hofmann-Winkler H, Thiel V, Pohlmann S.** 2013. TMPRSS2 activates the human coronavirus 229E for cathepsin-independent host cell entry and is expressed in viral target cells in the respiratory epithelium. J Virol **87:**6150-6160.
20. **Glowacka I, Bertram S, Muller MA, Allen P, Soilleux E, Pfefferle S, Steffen I, Tsegaye TS, He Y, Gnirss K, Niemeyer D, Schneider H, Drosten C, Pohlmann S.** 2011. Evidence that TMPRSS2 activates the severe acute respiratory syndrome coronavirus spike protein for membrane fusion and reduces viral control by the humoral immune response. J Virol **85:**4122-4134.
21. **Gallaher WR.** 1996. Similar structural models of the transmembrane proteins of Ebola and avian sarcoma viruses. Cell **85:**477-478.
22. **Hernandez LD, White JM.** 1998. Mutational analysis of the candidate internal fusion peptide of the avian leukosis and sarcoma virus subgroup A envelope glycoprotein. J Virol **72:**3259-3267.
23. **Delos SE, Gilbert JM, White JM.** 2000. The central proline of an internal viral fusion peptide serves two important roles. J Virol **74:**1686-1693.
24. **Sanchez A, Yang ZY, Xu L, Nabel GJ, Crews T, Peters CJ.** 1998. Biochemical analysis of the secreted and virion glycoproteins of Ebola virus. J Virol **72:**6442-6447.
25. **Cross KJ, Langley WA, Russell RJ, Skehel JJ, Steinhauer DA.** 2009. Composition and functions of the influenza fusion peptide. Protein Pept Lett **16:**766-778.
26. **Gallaher WR.** 1987. Detection of a fusion peptide sequence in the transmembrane protein of human immunodeficiency virus. Cell **50:**327-328.
27. **Martin I, Ruysschaert JM.** 2000. Common properties of fusion peptides from diverse systems. Bioscience reports **20:**483-500.
28. **Guillen J, de Almeida RF, Prieto M, Villalain J.** 2008. Structural and dynamic characterization of the interaction of the putative fusion peptide of the S2 SARS-CoV virus protein with lipid membranes. The journal of physical chemistry. B **112:**6997-7007.
29. **Guillen J, Kinnunen PK, Villalain J.** 2008. Membrane insertion of the three main membranotropic sequences from SARS-CoV S2 glycoprotein. Biochimica et biophysica acta **1778:**2765-2774.
30. **Guillen J, Perez-Berna AJ, Moreno MR, Villalain J.** 2005. Identification of the membrane-active regions of the severe acute respiratory syndrome coronavirus spike membrane glycoprotein using a 16/18-mer peptide scan: implications for the viral fusion mechanism. J Virol **79:**1743-1752.

31. **Guillen J, Perez-Berna AJ, Moreno MR, Villalain J.** 2008. A second SARS-CoV S2 glycoprotein internal membrane-active peptide. Biophysical characterization and membrane interaction. Biochemistry **47:**8214-8224.
32. **Madu IG, Roth SL, Belouzard S, Whittaker GR.** 2009. Characterization of a highly conserved domain within the severe acute respiratory syndrome coronavirus spike protein S2 domain with characteristics of a viral fusion peptide. J Virol **83:**7411-7421.
33. **Sainz B, Jr., Rausch JM, Gallaher WR, Garry RF, Wimley WC.** 2005. Identification and characterization of the putative fusion peptide of the severe acute respiratory syndrome-associated coronavirus spike protein. J Virol **79:**7195-7206.
34. **Luo Z, Weiss SR.** 1998. Roles in cell-to-cell fusion of two conserved hydrophobic regions in the murine coronavirus spike protein. Virology **244:**483-494.
35. **Chambers P, Pringle CR, Easton AJ.** 1990. Heptad repeat sequences are located adjacent to hydrophobic regions in several types of virus fusion glycoproteins. J Gen Virol **71 (Pt 12):**3075-3080.
36. **Jeffers SA, Tusell SM, Gillim-Ross L, Hemmila EM, Achenbach JE, Babcock GJ, Thomas WD, Jr., Thackray LB, Young MD, Mason RJ, Ambrosino DM, Wentworth DE, Demartini JC, Holmes KV.** 2004. CD209L (L-SIGN) is a receptor for severe acute respiratory syndrome coronavirus. Proc Natl Acad Sci U S A **101:**15748-15753.
37. **Qian Z, Dominguez SR, Holmes KV.** 2013. Role of the spike glycoprotein of human Middle East respiratory syndrome coronavirus (MERS-CoV) in virus entry and syncytia formation. PLoS One **8:**e76469.
38. **Peng G, Sun D, Rajashankar KR, Qian Z, Holmes KV, Li F.** 2011. Crystal structure of mouse coronavirus receptor-binding domain complexed with its murine receptor. Proc Natl Acad Sci U S A **108:**10696-10701.
39. **Zelus BD, Schickli JH, Blau DM, Weiss SR, Holmes KV.** 2003. Conformational changes in the spike glycoprotein of murine coronavirus are induced at 37 degrees C either by soluble murine CEACAM1 receptors or by pH 8. J Virol **77:**830-840.
40. **Qian Z, Wang H, Empig C, Anderson WF, Albritton LM.** 2004. Complementation of a binding-defective retrovirus by a host cell receptor mutant. J Virol **78:**5766-5772.
41. **Martin I, Schaal H, Scheid A, Ruysschaert JM.** 1996. Lipid membrane fusion induced by the human immunodeficiency virus type 1 gp41 N-terminal extremity is determined by its orientation in the lipid bilayer. J Virol **70:**298-304.
42. **Chen YH, Yang JT, Chau KH.** 1974. Determination of the helix and beta form of proteins in aqueous solution by circular dichroism. Biochemistry **13:**3350-3359.
43. **Hope MJ, Bally MB, Webb G, Cullis PR.** 1985. Production of large unilamellar vesicles by a rapid extrusion procedure: characterization of size distribution, trapped volume and ability to maintain a membrane potential. Biochimica et biophysica acta **812:**55-65.
44. **Struck DK, Hoekstra D, Pagano RE.** 1981. Use of resonance energy transfer to monitor membrane fusion. Biochemistry **20:**4093-4099.
45. **Gierer S, Bertram S, Kaup F, Wrensch F, Heurich A, Kramer-Kuhl A, Welsch K, Winkler M, Meyer B, Drosten C, Dittmer U, von Hahn T, Simmons G, Hofmann H, Pohlmann S.** 2013. The Spike Protein of the Emerging Betacoronavirus EMC Uses a Novel Coronavirus Receptor for Entry, Can Be Activated by TMPRSS2, and Is Targeted by Neutralizing Antibodies. J Virol **87:**5502-5511.

46. **Waring AJ, Mobley PW, Gordon LM.** 1998. Conformational mapping of a viral fusion peptide in structure-promoting solvents using circular dichroism and electrospray mass spectrometry. Proteins **Suppl 2:**38-49.
47. **Nelson JW, Kallenbach NR.** 1986. Stabilization of the ribonuclease S-peptide alpha-helix by trifluoroethanol. Proteins **1:**211-217.
48. **Gao J, Lu G, Qi J, Li Y, Wu Y, Deng Y, Geng H, Li H, Wang Q, Xiao H, Tan W, Yan J, Gao GF.** 2013. Structure of the fusion core and inhibition of fusion by a heptad repeat peptide derived from the S protein of Middle East respiratory syndrome coronavirus. J Virol **87:**13134-13140.
49. **Lu L, Liu Q, Zhu Y, Chan KH, Qin L, Li Y, Wang Q, Chan JF, Du L, Yu F, Ma C, Ye S, Yuen KY, Zhang R, Jiang S.** 2014. Structure-based discovery of Middle East respiratory syndrome coronavirus fusion inhibitor. Nature communications **5:**3067.
50. **Bosch BJ, Martina BE, Van Der Zee R, Lepault J, Haijema BJ, Versluis C, Heck AJ, De Groot R, Osterhaus AD, Rottier PJ.** 2004. Severe acute respiratory syndrome coronavirus (SARS-CoV) infection inhibition using spike protein heptad repeat-derived peptides. Proc Natl Acad Sci U S A **101:**8455-8460.
51. **Xu Y, Liu Y, Lou Z, Qin L, Li X, Bai Z, Pang H, Tien P, Gao GF, Rao Z.** 2004. Structural basis for coronavirus-mediated membrane fusion. Crystal structure of mouse hepatitis virus spike protein fusion core. J Biol Chem **279:**30514-30522.
52. **Xu Y, Lou Z, Liu Y, Pang H, Tien P, Gao GF, Rao Z.** 2004. Crystal structure of severe acute respiratory syndrome coronavirus spike protein fusion core. J Biol Chem **279:**49414-49419.
53. **Bosch BJ, van der Zee R, de Haan CA, Rottier PJ.** 2003. The coronavirus spike protein is a class I virus fusion protein: structural and functional characterization of the fusion core complex. J Virol **77:**8801-8811.
54. **Kaufman G, Liu P, Leibowitz JL.** 2014. Identification of novel functional regions within the spike glycoprotein of MHV-A59 based on a bioinformatics approach. Virus Res **189:**177-188.
55. **McRoy WC, Baric RS.** 2008. Amino acid substitutions in the S2 subunit of mouse hepatitis virus variant V51 encode determinants of host range expansion. J Virol **82:**1414-1424.
56. **Liao Y, Zhang SM, Neo TL, Tam JP.** 2015. Tryptophan-dependent membrane interaction and heteromerization with the internal fusion peptide by the membrane proximal external region of SARS-CoV spike protein. Biochemistry **54:**1819-1830.
57. **Li L, Vorobyov I, MacKerell AD, Jr., Allen TW.** 2008. Is arginine charged in a membrane? Biophysical journal **94:**L11-13.
58. **Lai AL, Park H, White JM, Tamm LK.** 2006. Fusion peptide of influenza hemagglutinin requires a fixed angle boomerang structure for activity. J Biol Chem **281:**5760-5770.
59. **Li J, Das P, Zhou R.** 2010. Single mutation effects on conformational change and membrane deformation of influenza hemagglutinin fusion peptides. The journal of physical chemistry. B **114:**8799-8806.
60. **Li Y, Han X, Lai AL, Bushweller JH, Cafiso DS, Tamm LK.** 2005. Membrane structures of the hemifusion-inducing fusion peptide mutant G1S and the fusion-blocking mutant G1V of influenza virus hemagglutinin suggest a mechanism for pore opening in membrane fusion. J Virol **79:**12065-12076.

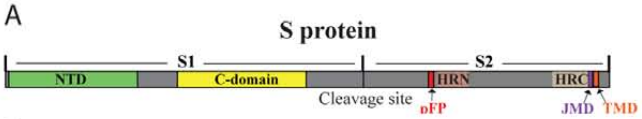
61. **Lai AL, Freed JH.** 2014. HIV gp41 fusion peptide increases membrane ordering in a cholesterol-dependent fashion. Biophysical journal **106:**172-181.
62. **Fuhrmans M, Marrink SJ.** 2012. Molecular view of the role of fusion peptides in promoting positive membrane curvature. Journal of the American Chemical Society **134:**1543-1552.
63. **Gregory SM, Harada E, Liang B, Delos SE, White JM, Tamm LK.** 2011. Structure and function of the complete internal fusion loop from Ebolavirus glycoprotein 2. Proc Natl Acad Sci U S A **108:**11211-11216.
64. **Lau WL, Ege DS, Lear JD, Hammer DA, DeGrado WF.** 2004. Oligomerization of fusogenic peptides promotes membrane fusion by enhancing membrane destabilization. Biophysical journal **86:**272-284.
65. **Yang J, Prorok M, Castellino FJ, Weliky DP.** 2004. Oligomeric beta-structure of the membrane-bound HIV-1 fusion peptide formed from soluble monomers. Biophysical journal **87:**1951-1963.
66. **Chang DK, Cheng SF, Kantchev EA, Lin CH, Liu YT.** 2008. Membrane interaction and structure of the transmembrane domain of influenza hemagglutinin and its fusion peptide complex. BMC biology **6:**2.
67. **Donald JE, Zhang Y, Fiorin G, Carnevale V, Slochower DR, Gai F, Klein ML, DeGrado WF.** 2011. Transmembrane orientation and possible role of the fusogenic peptide from parainfluenza virus 5 (PIV5) in promoting fusion. Proc Natl Acad Sci U S A **108:**3958-3963.
68. **Reuven EM, Dadon Y, Viard M, Manukovsky N, Blumenthal R, Shai Y.** 2012. HIV-1 gp41 transmembrane domain interacts with the fusion peptide: implication in lipid mixing and inhibition of virus-cell fusion. Biochemistry **51:**2867-2878.
69. **Teese MG, Langosch D.** 2015. Role of GxxxG Motifs in Transmembrane Domain Interactions. Biochemistry **54:**5125-5135.
70. **Faingold O, Cohen T, Shai Y.** 2012. A GxxxG-like motif within HIV-1 fusion peptide is critical to its immunosuppressant activity, structure, and interaction with the transmembrane domain of the T-cell receptor. J Biol Chem **287:**33503-33511.
71. **Chang DK, Cheng SF, Chien WJ.** 1997. The amino-terminal fusion domain peptide of human immunodeficiency virus type 1 gp41 inserts into the sodium dodecyl sulfate micelle primarily as a helix with a conserved glycine at the micelle-water interface. J Virol **71:**6593-6602.
72. **Sackett K, Shai Y.** 2005. The HIV fusion peptide adopts intermolecular parallel beta-sheet structure in membranes when stabilized by the adjacent N-terminal heptad repeat: a 13C FTIR study. Journal of molecular biology **350:**790-805.
73. **Lorieau JL, Louis JM, Bax A.** 2010. The complete influenza hemagglutinin fusion domain adopts a tight helical hairpin arrangement at the lipid:water interface. Proc Natl Acad Sci U S A **107:**11341-11346.
74. **Haria NR, Monticelli L, Fraternali F, Lorenz CD.** 2014. Plasticity and conformational equilibria of influenza fusion peptides in model lipid bilayers. Biochimica et biophysica acta **1838:**1169-1179.

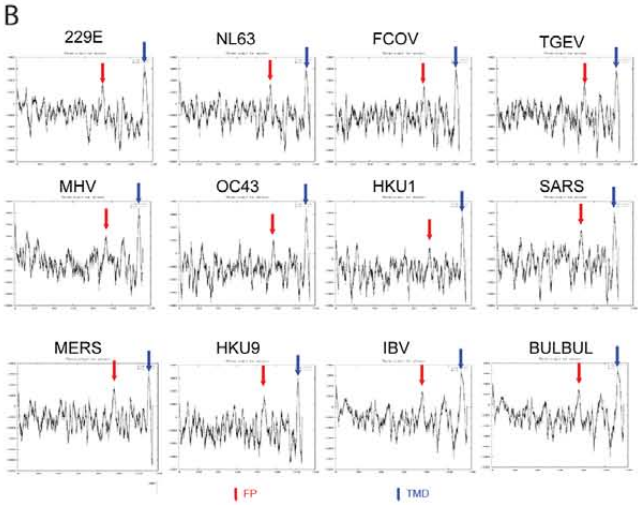
Table I Summary of pFP mutants of betacoronaviruses

	Expression in cell lysate	Expression on cell surface	Incorporation in viron	Receptor binding	Cell-cell fusion	Pseudovirion transduction
MERS wild type	+++	+++	+++	+++	+++	+++
S949G	+++	+++	+++	+++	+++	+++
S950G	+++	+++	+++	+++	+++	+++
L951G	+	-	-	-	-	-
L952G	+	-	-	-	-	-
L952A	++	+++	++	+++	++	+
L951G/L952G	+	-	-	-	-	-
G953A	+++	+++	++	+++	+++	+
G953R	+++	+++	++	+++	-	-
S954G	+++	+++	+++	+++	+++	+++
S954R	+++	+++	+++	+++	-	-
I955G	+++	+++	+++	+++	-	-
A956V	+++	+++	+++	+++	++	+++
A956R	+++	+++	+++	+++	+++	+++
G957A	+++	+++	+++	+++	+++	++
G957R	+++	+++	+++	+++	-	-
V958G	+++	+++	+++	+++	+++	+++
V958R	+++	+++	+++	+++	-	-
I955G/V958G	+++	+++	+++	+++	-	-
G959A	+++	+++	+++	+++	+++	+++
G959R	+++	+++	+++	+++	+	-
W960G	+++	+++	+++	+++	+	+
W960R	+++	+++	+++	+++	-	-
V958G/W960G	+++	+++	+++	+++	-	-
T961G	+++	+++	+++	+++	-	-
A962V	+++	+++	+++	+++	+++	+++
A962R	+++	+++	+++	+++	+	-
G963A	+++	+++	+++	+++	++	++
G963R	+++	+++	++	+++	-	-
L964G	+++	+++	+++	+++	+++	++
L964R	+++	+++	+++	+++	-	-
S965G	+++	+++	+++	+++	+++	+++
S965R	+	+	+	+	-	-
S966G	++	+++	+	++	+++	+
S966R	+	-	-	+	-	-
F967G	+	+	-	+	-	-
L964G/F967G	+	+	-	+	-	-
A968V	+++	+++	+++	++	+++	++
A968R	+	+	-	+	-	-
A969V	+	+	-	+	+	-
A969R	+	+	+	+	-	-
I970G	+	+	+	+	-	-
P971V	+	+	-	-	-	-
F972G	+	+	-	-	-	-
I970G/F972G	+	+	-	-	-	-
SARS wild type	+++	+++	+++	+++	+++	+++
W868R	+++	**ND**	+++	**ND**	+	-
W868G	+++	+++	+++	+++	++	+++
F870R	+++	**ND**	+++	**ND**	+	-
F870G	+++	+++	+++	+++	++	+++
W868G/F870G	+++	+++	+++	+++	-	-
L876R	+++	**ND**	+++	**ND**	-	-
L876G	+++	+++	+++	+++	++	+
I878R	++	**ND**	++	**ND**	-	-
I878G	++	+++	++	+++	+	-
L876G/I878G	++	+++	++	+++	-	-
MHV wild type	+++	+++	+++	+++	+++	+++
M963R	+	+	++	++	+	-
F937R	+++	+++	+++	+++	-	-
P938R	++	++	+++	+++	-	-
P939R	+++	+++	+++	+++	+	-
W940R	+++	+++	+++	+++	-	-

Western blot analysis on S protein expression in cell lysate : +++, very strong; ++, strong; +, weak; -, absent
Western blot analysis on S protein expression incorporation in virion: +++, strong; ++, weak; +, very weak; -, absent

S protein expression on cell surface and receptor binding: +++, >70% of WT; ++, 46-70% of WT; +, 20-45% of WT; -, <20% of WT. ND, not done.
Cell-cell fusion and pseudovirion transduction: +++, >70% of WT; ++, 31-70% of WT; +, 5-30% of WT; -, <5% of WT. ND, not done.





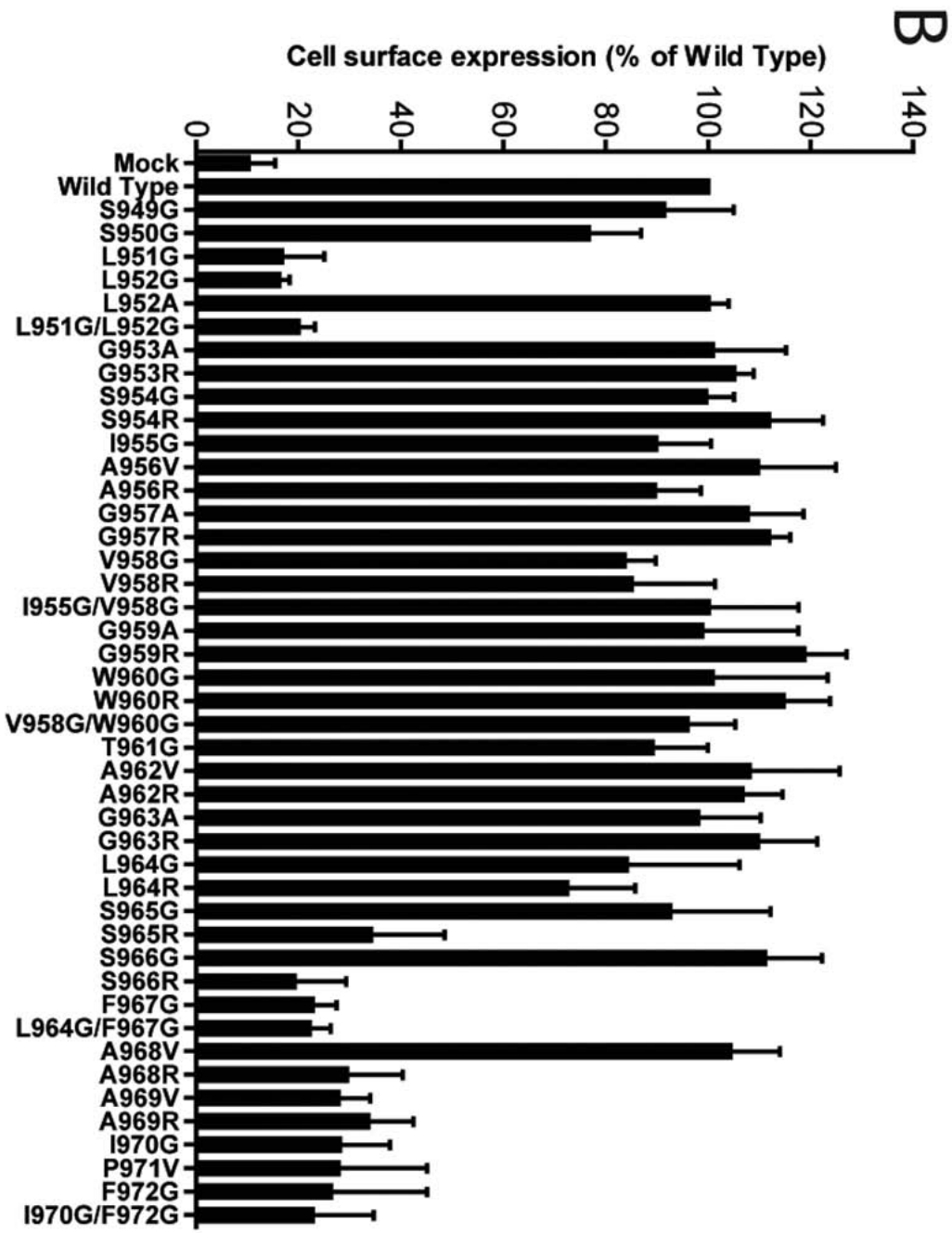
C

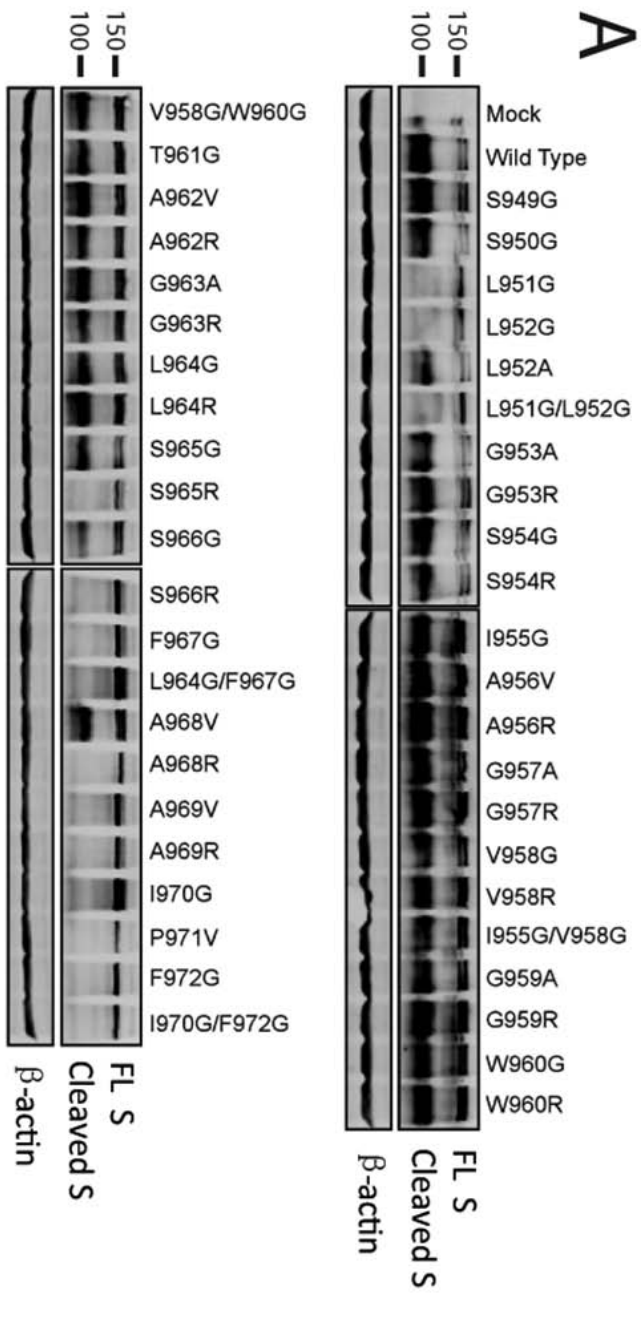
CoV				Group
229E	748	YTGSLIGGIALG----GLTS-AVSIPF	769	alpha
NL63	929	YTGSLIGGMVLG----GLTS-AAAIPF	950	alpha
FCoV	1035	YTASLIGGMALG----SITS-AVAVPF	1056	alpha
TGEV	1017	YTASLAGGITLG----ALGGGAVAIPF	1039	alpha
MHV	927	YTTGATAAAMFP----PWSA-AAGVPF	948	beta A
BCoV	971	YTLAATSASLFP----PLSA-AVGVPF	992	beta A
OC43	961	YTLAATSASLFP----PWTA-AAGVPF	982	beta A
HKU1	958	YTTAATVAAMFP----PWSA-AAGIPF	979	beta A
SARS	855	YTAALVSGTATAGWTFGAGA-ALQIPF	880	beta B
MERS	947	YTSSLLGSIAGVGWTAGLSS-FAAIPF	972	beta C
HKU4	946	YTSSLLGSIAGAGWTAGLSS-FAAIPF	971	beta C
HKU9	870	YTSLLVGAVASSGYTFGITS-AGVIPF	895	beta D
IBV	757	YTSSLVASMAFG----GITA-AGAIPF	778	gamma
Bulbul	737	YTASLTGAMVFG----GMTA-AAAIPF	758	delta

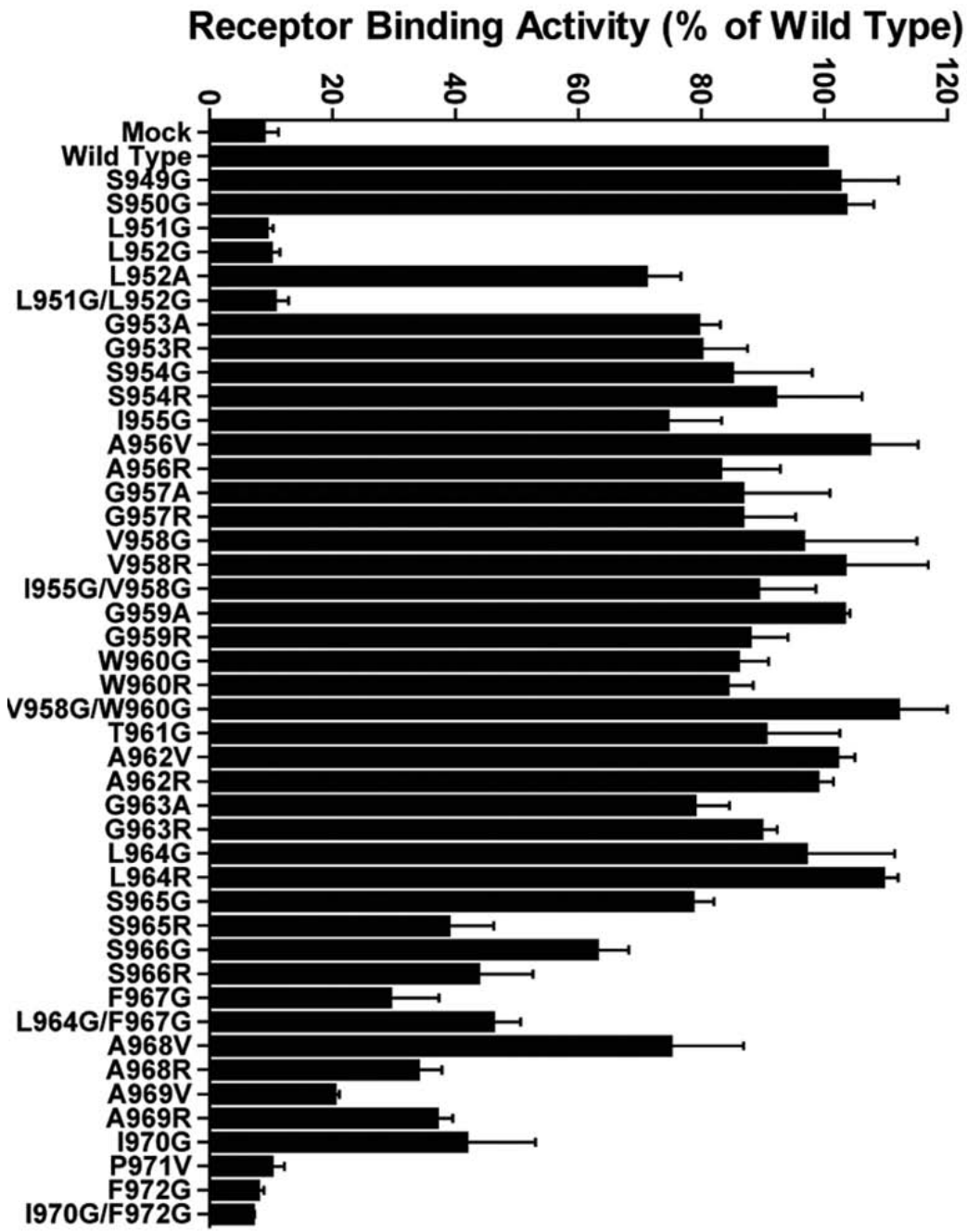
D

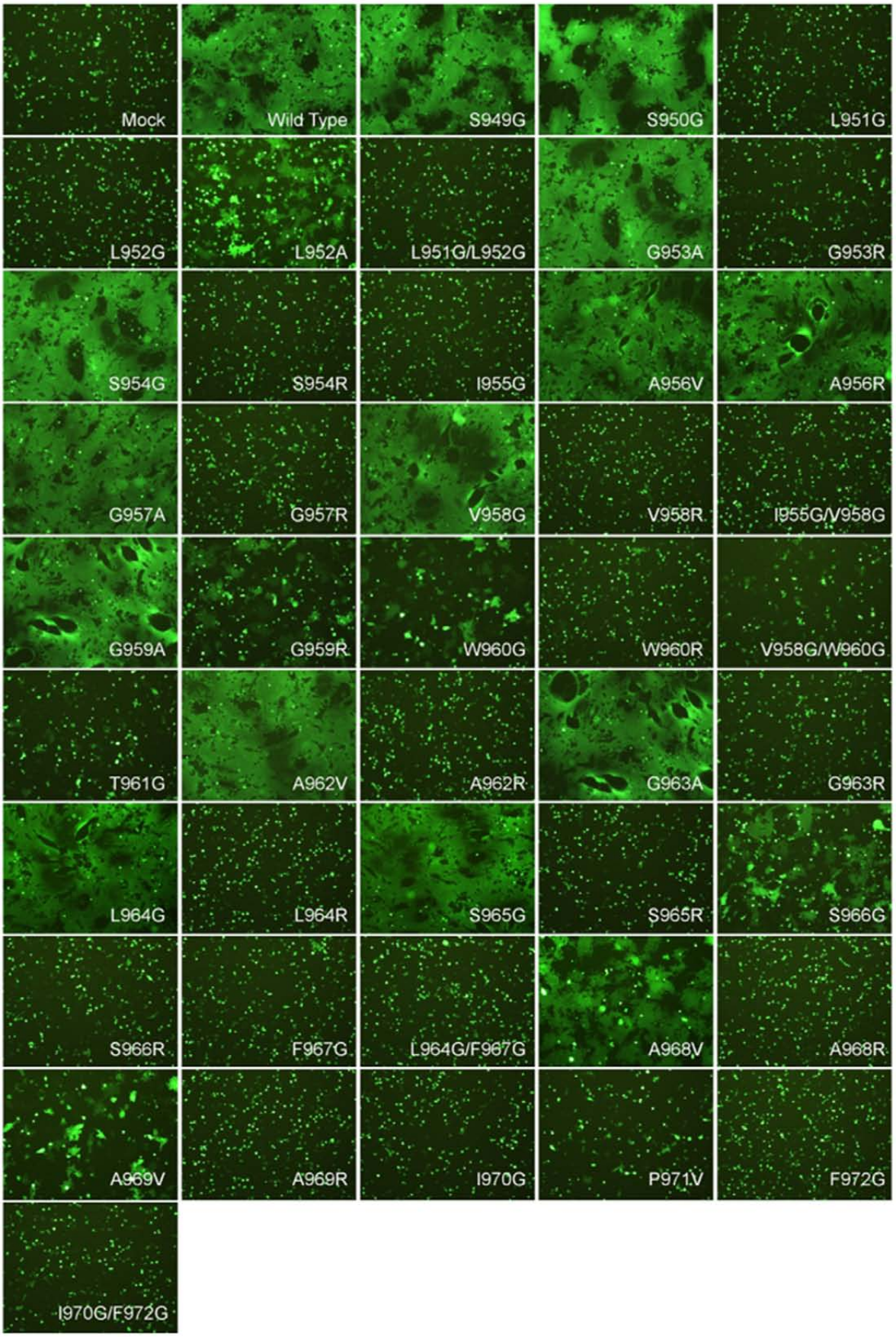
MERS	S	S	L	L	G	S	I	A	G	V	G	W	T	A	G	L	S	S	F	A	A	I	P	F
S949G	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
S950G	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
L951G	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
L952G	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
L952A	-	-	-	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
L951G/L952G	-	-	G	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
G953A	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
G953R	-	-	-	-	R	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
S954G	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
S954R	-	-	-	-	-	R	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
I955G	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
A956V	-	-	-	-	-	-	-	V	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
A956R	-	-	-	-	-	-	-	R	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
G957A	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
G957R	-	-	-	-	-	-	-	-	R	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
V958G	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-
V958R	-	-	-	-	-	-	-	-	-	R	-	-	-	-	-	-	-	-	-	-	-	-	-	-
I955G/V958G	-	-	-	-	-	-	G	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-
G959A	-	-	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-	-	-	-
G959R	-	-	-	-	-	-	-	-	-	-	R	-	-	-	-	-	-	-	-	-	-	-	-	-
W960G	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-
W960R	-	-	-	-	-	-	-	-	-	-	-	R	-	-	-	-	-	-	-	-	-	-	-	-
V958G/W960G	-	-	-	-	-	-	-	-	-	G	-	G	-	-	-	-	-	-	-	-	-	-	-	-
T961G	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-
A962V	-	-	-	-	-	-	-	-	-	-	-	-	-	V	-	-	-	-	-	-	-	-	-	-
A962R	-	-	-	-	-	-	-	-	-	-	-	-	-	R	-	-	-	-	-	-	-	-	-	-
G963A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-
G963R	-	-	-	-	-	-	-	-	-	-	-	-	-	-	R	-	-	-	-	-	-	-	-	-
L964G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-
L964R	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	R	-	-	-	-	-	-	-	-
S965G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-
S965R	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	R	-	-	-	-	-	-	-
S966G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-
S966R	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	R	-	-	-	-	-	-
F967G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-
L964G/F967G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	G	-	-	-	-	-
A968V	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	V	-	-	-	-
A968R	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	R	-	-	-	-
A969V	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	V	-	-	-
A969R	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	R	-	-	-
I970G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-
P971V	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-
F972G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G
I970G/F972G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	G

Journal of Virology


B
Cell surface expression (% of Wild Type)
0
20
40
60
80
100
120
140
Mock
Wild Type
S949G
S950G
L951G
L952G
L952A
L951G/L952G
G953A
G953R
S954G
S954R
I955G
A956V
A956R
G957A
G957R
V958G
V958R
I955G/V958G
G959A
G959R
W960G
W960R
V958G/W960G
T961G
A962V
A962R
G963A
G963R
L964G
L964R
S965G
S965R
S966G
S966R
F967G
L964G/F967G
A968V
A968R
A969V
A969R
I970G
P971V
F972G
I970G/F972G

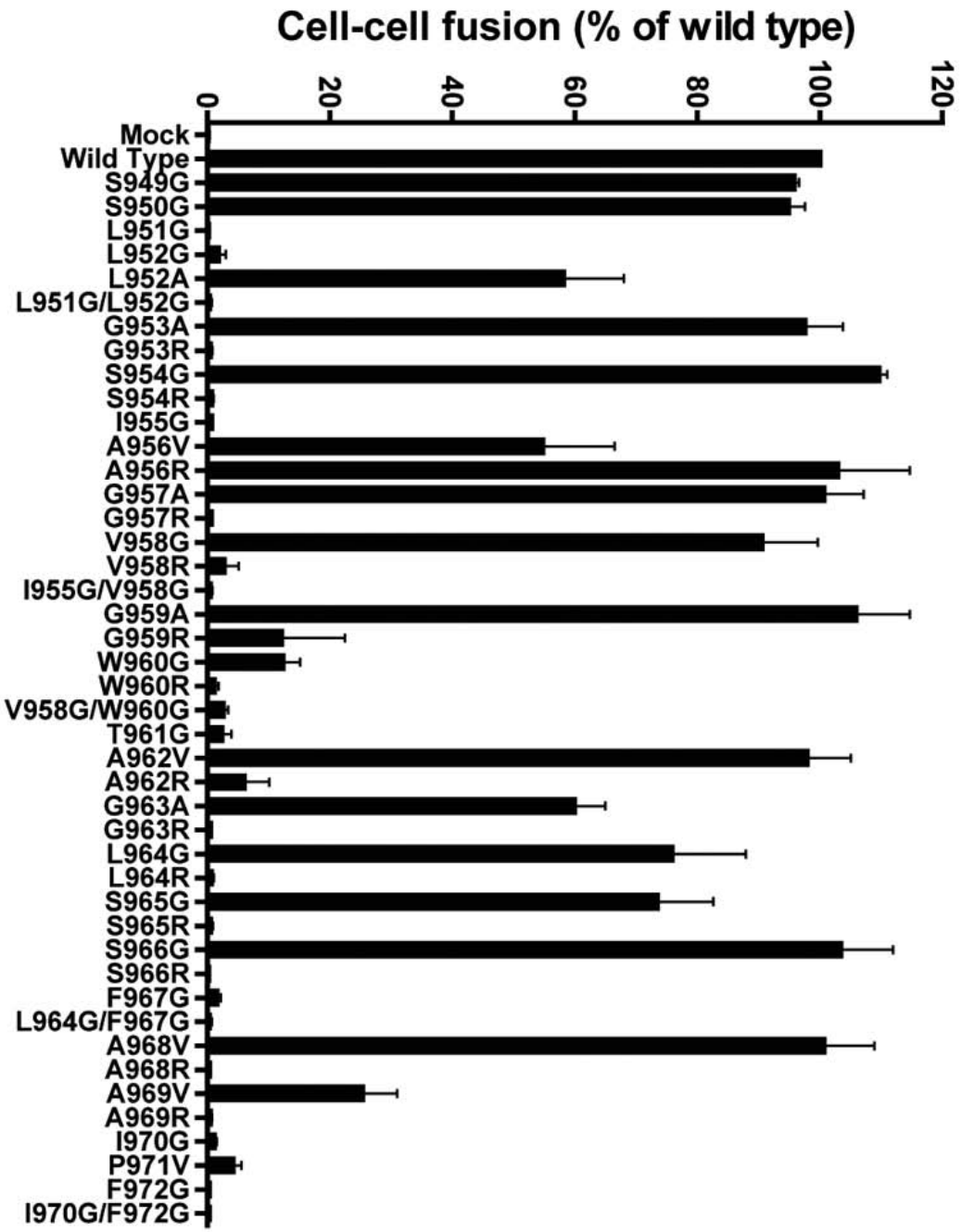

A
100
150
Mock
Wild Type
S949G
S950G
L951G
L952G
L952A
L951G/L952G
G953A
G953R
S954G
S954R
I955G
A956V
A956R
G957A
G957R
V958G
V958R
I955G/V958G
G959A
G959R
W960G
W960R
FL S
Cleaved S
β-actin
100
150
V958G/W960G
T961G
A962V
A962R
G963A
G963R
L964G
L964R
S965G
S965R
S966G
S966R
F967G
L964G/F967G
A968V
A968R
A969V
A969R
I970G
P971V
F972G
I970G/F972G
FL S
Cleaved S
β-actin

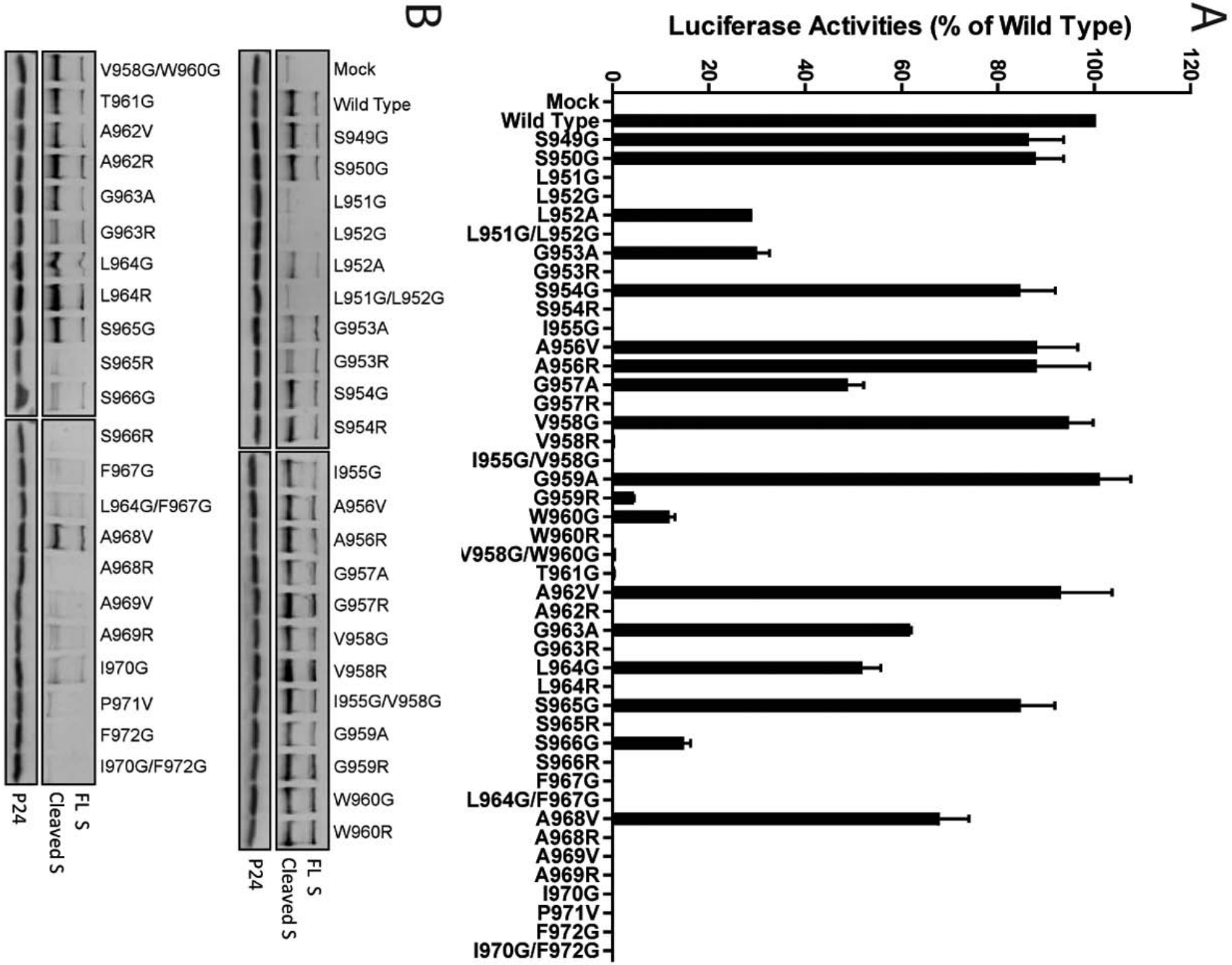

Receptor Binding Activity (% of Wild Type)
0
20
40
60
80
100
120
Mock
Wild Type
S949G
S950G
L951G
L952G
L952A
L951G/L952G
G953A
G953R
S954G
S954R
I955G
A956V
A956R
G957A
G957R
V958G
V958R
I955G/V958G
G959A
G959R
W960G
W960R
V958G/W960G
T961G
A962V
A962R
G963A
G963R
L964G
L964R
S965G
S965R
S966G
S966R
F967G
L964G/F967G
A968V
A968R
A969V
A969R
I970G
P971V
F972G
I970G/F972G

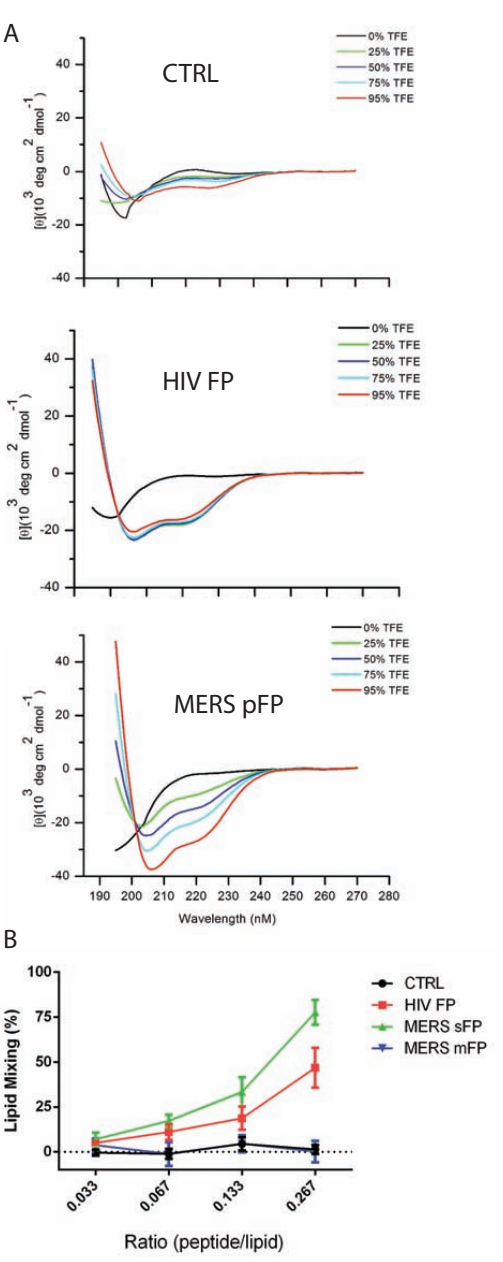


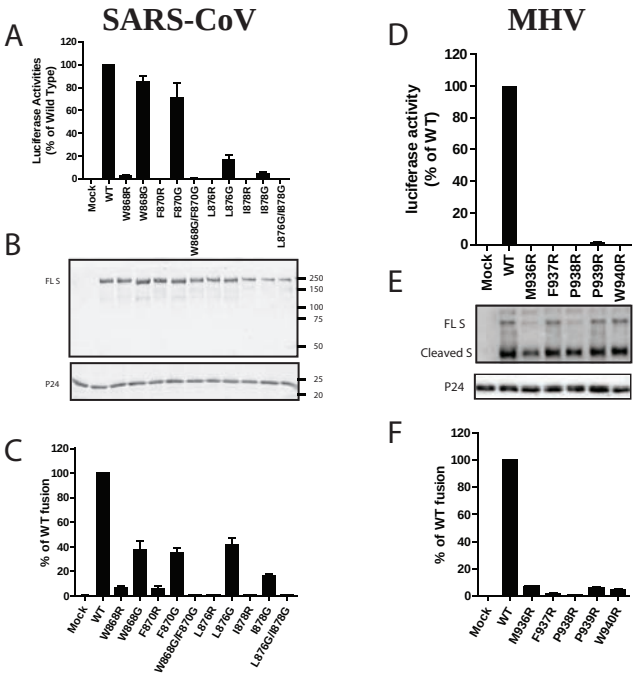
JVI Journal of Virology

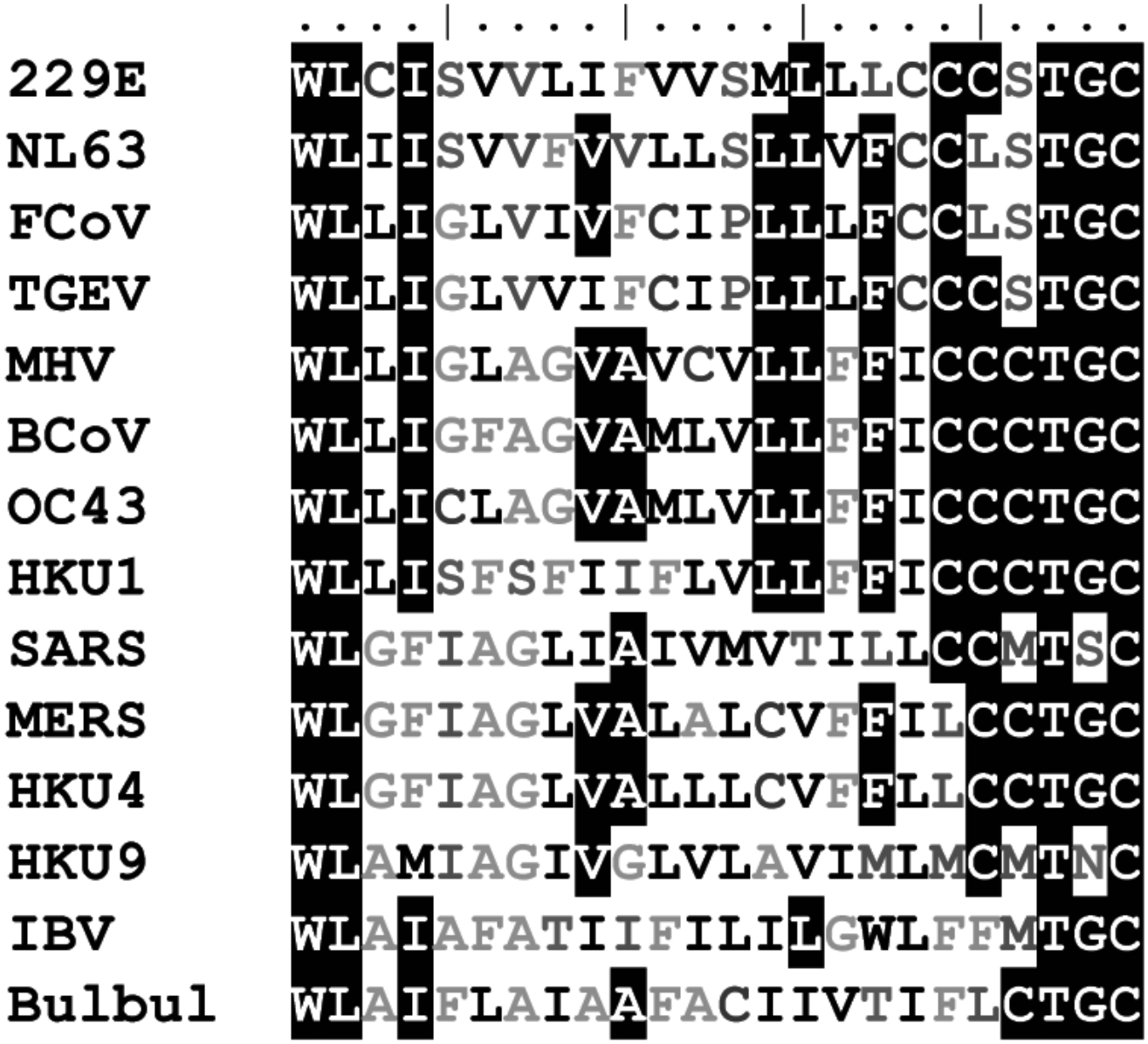
Journal of Virology


Cell-cell fusion (% of wild type)
0
20
40
60
80
100
120
Mock
Wild Type
S949G
S950G
L951G
L952G
L952A
L951G/L952G
G953A
G953R
S954G
S954R
I955G
A956V
A956R
G957A
G957R
V958G
V958R
I955G/V958G
G959A
G959R
W960G
W960R
V958G/W960G
T961G
A962V
A962R
G963A
G963R
L964G
L964R
S965G
S965R
S966G
S966R
F967G
L964G/F967G
A968V
A968R
A969V
A969R
I970G
P971V
F972G
I970G/F972G


A
Luciferase Activities (% of Wild Type)
0
20
40
60
80
100
120
Mock
Wild Type
S949G
S950G
L951G
L952G
L952A
L951G/L952G
G953A
G953R
S954G
S954R
I955G
A956V
A956R
G957A
G957R
V958G
V958R
I955G/V958G
G959A
G959R
W960G
W960R
V958G/W960G
T961G
A962V
A962R
G963A
G963R
L964G
L964R
S965G
S965R
S966G
S966R
F967G
L964G/F967G
A968V
A968R
A969V
A969R
I970G
P971V
F972G
I970G/F972G
B
V958G/W960G
T961G
A962V
A962R
G963A
G963R
L964G
L964R
S965G
S965R
S966G
S966R
F967G
L964G/F967G
A968V
A968R
A969V
A969R
I970G
P971V
F972G
I970G/F972G
P24
Cleaved S
FL S
Mock
Wild Type
S949G
S950G
L951G
L952G
L952A
L951G/L952G
G953A
G953R
S954G
S954R
I955G
A956V
A956R
G957A
G957R
V958G
V958R
I955G/V958G
G959A
G959R
W960G
W960R
P24
Cleaved S
FL S







Journal of Virology