

MESSENGER RNA VACCINES AGAINST WIDE SPECTRUM OF CORONAVIRUS VARIANTS

SEQUENCE LISTING

[0001] The instant application contains a Sequence Listing which is submitted electronically in ASCII format and is hereby incorporated by reference in its entirety. The ASCII copy, created on April 12, 2022, is named "G4590-15000PCT_SeqListing_20220412" and is 111 kilobytes in size.

CROSS-REFERENCE TO RELATED APPLICATIONS

[0002] This application claims benefit to and priority to U.S. Provisional Patent Application No. 63/173,752, filed on April 12, 2021, and U.S. Provisional Patent Application No. 63/264,737 filed on December 1, 2021, the contents of which are hereby incorporated by reference in their entirety.

FIELD

[0003] The present disclosure relates generally to the field of treating and/or preventing a coronavirus infection. Particularly, the present disclosure relates to messenger RNA (mRNA) vaccines against wide spectrum of coronavirus (CoV) variants.

BACKGROUND OF THE INVENTION

[0004] In 1796, Edward Jenner created the first vaccine (cowpox) in the world to protect against smallpox and successfully rescued millions of people. Since then, vaccination has been recognized as the best way to protect against pathogens. Since the outbreak of severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) in December 2019 that caused Coronavirus Induced Disease 2019 (COVID-19), the virus has spread all over the world and caused more than 200 million infections and 4 million deaths in 20 months. This pandemic has become a major threat to public health.

[0005] A great number of efforts has been directed toward the development of effective means and vaccines to combat this pandemic. The trimeric spike (S) protein on the viral surface has been the key immunogen and target for preventive vaccine and therapeutic antibody development. As of December 2020, the U.S. FDA has authorized Pfizer/BioNTech and Moderna's mRNA vaccine candidates as well as Regeneron's antibodies, for emergency use; however, several other vaccine candidates and human antibodies are in clinical trials and a handful of them, including the Oxford/Astrazeneca and J&J vaccines, are close to obtaining approval. Of the various vaccines developed to control the spread of SARS-CoV-2 and variants, the mRNA vaccines developed by Moderna and BioNTech/Pfizer represent a major breakthrough due to their speed and convenience.

These vaccines were stabilized with new mRNA technology and lipid nanoparticle (LNP) formulation for delivery and translation into the spike (S) protein *in vivo* to induce immune response (*Ewen Callaway. COVID vaccine excitement builds as Moderna reports third positive result. Nature. 587 (7834):337-338 (2020); Polack P. et al. Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine. N Engl J Med. 383 (27):2603-2615 (2020).*

[0006] However, since there have been a number of variants that have been circulating widely in the world and more infectious variants such as beta, delta and omicron variants may have an increased ability to re-infect people who have been vaccinated previously and/or recovered from infection by earlier versions of the coronavirus, infections by emerging SARS-CoV-2 variants may continue to occur or potentially increase in frequency. The S protein of this RNA virus is highly glycosylated and frequently mutated with more than 9 million sequences and over 1,000 sites of mutation in its 1,273 amino-acid sequence reported to GISAID (www.gisaid.org), including the highly transmissible delta and omicron variants, posing a major challenge in the development of broadly effective antibodies and vaccines.

[0007] Accordingly, there exists an immediate need for better vaccines as well as better products and methods to prevent and treat coronavirus infections.

SUMMARY OF THE INVENTION

[0008] The present disclosure provides a novel coronavirus mRNA vaccine, methods of preparation and uses thereof. The novel vaccine is designed based on an mRNA technology to remove the glycan shields of a coronavirus (*e.g.* SARS-CoV-2) spike protein to better expose the conserved regions of the spike protein. The mRNA vaccine of coronavirus spike protein has deletion of glycosites in the receptor binding domain (RBD) or the subunit 2 (S2) domain to expose highly conserved epitopes and elicit antibodies and CD8 T-cell response with broader protection against the alpha, beta, gamma, delta, omicron and various variants, as compared to the unmodified mRNA. The mRNA vaccine provided herein is effective for inducing protective immunity against SARS-CoV-2 and variants (*e.g.* alpha, beta, gamma, delta, omicron). When used individually or in combination as an immunogenic composition or vaccine, the mRNA vaccine of the present disclosure may protect people from infection and/or to reduce symptoms if infected.

[0009] In one aspect, the present disclosure provides at least one immunogenic peptide, comprising an amino acid sequence selected from a group consisting of: TESIVRFPNITNL (SEQ ID NO: 41), NITNLCPFGEVFNATR (SEQ ID NO: 42), LYNSASFSTFK (SEQ ID NO: 43), LDSKVGGNYN (SEQ ID NO: 44), KSNLKPFRDIST (SEQ ID NO: 45),

KPFERDISTEIQAG (SEQ ID NO: 46), GPKKSTNLVKNKC (SEQ ID NO: 47), NCDVVIGIVNNTVY (SEQ ID NO: 48), PELDSFKEELDKYFK[N]HTS (SEQ ID NO: 49), VNIQKEIDRLNEVA (SEQ ID NO: 50), NLNESLIDLQ (SEQ ID NO: 51) and LGKYEQYIKWP (SEQ ID NO: 52) or an amino acid sequence having at least about 99%, 98%, 97%, 96%, 95% or 90% identity to any of SEQ ID NOs: 41 to 52.

[0010] In some embodiments, the immunogenic peptide comprises at least an amino acid sequence selected from a group consisting of SEQ ID NOs: 41 to 43 and 45 to 51. In some embodiments, the immunogenic peptide comprises at least one, two, three, four, five, six, seven, eight, nine, ten, eleven or twelve amino acids of SEQ ID NOs: 41 to 52. In some embodiments, the immunogenic peptide comprises at least one, two, three, four, five, six, seven, eight, nine, ten amino acids of SEQ ID NOs: SEQ ID NOs: 41 to 43 and 45 to 51.

[0011] In one aspect, the present disclosure provides a modified nucleic acid molecule encoding a modified spike protein comprising one or more amino acid substitutions at N-linked glycosylation sequons (N-X-S/T), wherein X is any amino acid residue except proline, and S/T denotes a serine or threonine residue.

[0012] In some embodiments, the modified spike protein described herein comprises the the substitution of asparagine (N) to glutamine (Q) at N-linked glycosylation sequons (N-X-S/T) to eliminate N-linked glycan sequons.

[0013] In some embodiments, the modified spike protein described herein comprises one or more amino acid substitution at N-linked glycosylation sequons (N-X-S/T) to eliminate N-linked glycan sequons.

[0014] In some embodiments, the modified spike protein described herein comprises one or more amino acid substitutions of S/T at O-linked glycosylation sites to eliminate O-linked glycosylation sites. One example is the substitution of S/T to alanine (A).

[0015] In one embodiment, the modified nucleic acid molecule is an mRNA or a double-strand or single-strand DNA.

[0016] In one embodiment, the modified spike protein is derived from a SARS-CoV-2 spike protein. The SARS-CoV-2 spike protein described herein comprises an amino acid sequence of SEQ ID NO: 2, 16, 18 or 20, or amino acid sequence having at least about 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 2, 16, 18 or 20.

[0017] In some embodiments, the nucleic acid molecule encoding the amino acid sequence of SEQ ID NO: 2, 16, 18 or 20 is an mRNA comprising the nucleotide sequence of SEQ ID NO: 1,

15, 17 or 19 respectively, or a nucleotide sequence having at least about 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 1, 15, 17 or 19 respectively.

[0018] In some embodiments, the modified spike protein described herein comprises an amino acid sequence of SEQ ID NO: 4, 22, 24 or 26, wherein the modified spike protein comprises a receptor binding domain (RBD) lacking glycosylation sites. The modified nucleic acid molecule encoding the amino acid sequence of SEQ ID NO: 4, 22, 24 or 26 comprises the nucleotide sequence of SEQ ID NO: 3, 21, 23 or 25 respectively, or a nucleotide sequence having at least about 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 3, 21, 23 or 25 respectively.

[0019] In some embodiments, the modified spike protein described herein comprises an amino acid sequence of SEQ ID NO: 6, 28, 30 or 32, wherein the modified spike protein comprises a S2 subunit lacking glycosylation sites. The modified nucleic acid molecule encoding the amino acid sequence of SEQ ID NO: 6, 28, 30 or 32 comprises the nucleotide sequence of SEQ ID NO: 5, 27, 29 or 31 respectively, or a nucleotide sequence having at least about 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 5, 27, 29 or 31 respectively.

[0020] In some embodiments, the modified spike protein described herein comprises an amino acid sequence of SEQ ID NO: 8 or 34, wherein the modified spike protein comprises an S2 subunit that consists of a single glycosylation site. In some embodiments, the single glycosylation site is at the position N1194. In some embodiments, the modified nucleic acid molecule encoding the amino acid sequence of SEQ ID NO: 8 or 34 comprises the nucleotide sequence of SEQ ID NO: 7 or 33 respectively, or a nucleotide sequence having at least about 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 7 or 33 respectively.

[0021] In some embodiments, the modified spike protein described herein comprises an amino acid sequence of SEQ ID NO: 10 or 36, wherein the modified spike protein comprises a receptor binding domain (RBD) lacking glycosylation sites, and an amino acid substitution of N801 to Q801. The modified nucleic acid molecule encoding the amino acid sequence of SEQ ID NO: 10 or 36 comprises the nucleotide sequence of SEQ ID NO: 9 or 35 respectively, or a nucleotide sequence having at least about 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 9 or 35 respectively.

[0022] In some embodiments, the modified spike protein described herein comprises an amino acid sequence of SEQ ID NO: 12 or 38, wherein the modified spike protein comprises a receptor

binding domain (RBD) lacking glycosylation sites, and an amino acid substitution of N1194 to Q1194. The modified nucleic acid molecule encoding the amino acid sequence of SEQ ID NO: 12 or 38 comprises the nucleotide sequence of SEQ ID NO: 11 or 37 respectively, or a nucleotide sequence having at least about 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 11 or 37 respectively.

[0023] In some embodiments, the modified spike protein described herein comprises an amino acid sequence of SEQ ID NO: 14 or 40, wherein the modified spike protein comprises a modified receptor binding domain (RBD) lacking glycosylation sites, and amino acid substitutions of N122 to Q122, N165 to Q165, and N234 to Q234. The modified nucleic acid molecule encoding the amino acid sequence of SEQ ID NO: 14 or 40 comprises the nucleotide sequence of SEQ ID NO: 13 or 39 respectively, or a nucleotide sequence having at least about 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 13 or 39 respectively.

[0024] In some embodiments, the modified spike protein described herein comprises an S1 subunit lacking glycosylation sites.

[0025] In some embodiments, the modified spike protein described herein comprises both S1 and S2 subunits lacking glycosylation sites.

[0026] The present invention relates to the mRNA vaccine of coronavirus spike protein with deletion of glycosites in the receptor binding domain (RBD) or the subunit 2 (S2) domain to expose highly conserved epitopes and elicit antibodies and CD8 T-cell response with broader protection against the alpha, beta, gamma, delta, omicron and various variants, as compared to the unmodified mRNA.

[0027] In some embodiments, the coronavirus vaccine comprises a coronavirus spike protein mRNA with one or more mutations of the glycosites in RBD or S2 or other domains with one or more replacements of N to Q or S/T to A, or a combination thereof. In a further embodiment, the mutation of the N-glycosites is to change the putative sequon N-X-S/T to Q-X-S/T and/or change S/T of the O-glycosite to A.

[0028] In some embodiments, the mRNAs described herein having the glycosites with N to Q replacement include a S-(deg-RBD) (S protein with all 2 N-glycosites in RBD mutated from N to Q and 2 O-glycosites mutated from S/T to A), a S-(deg-S2) (S protein with all 9 glycosites in S2 mutated from N to Q), a S-(deg-S2-1194) (S protein with 8 glycosites in S2 mutated from N to Q, except glycosite 1194), a S-(deg-RBD-801) (S protein with all 2 N-glycosites in RBD mutated from N to Q and 2 O-glycosites mutated from S/T to A, and glycosite 801 mutated from N to Q), a S-(deg-RBD-1194) (S protein with all 2 N-glycosites in RBD mutated from N to Q and 2 O-

glycosites mutated from S/T to A, and glycosite 1194 mutated from N to Q), and a S-(deg-RBD-122-165-234) (S protein with all 2 N-glycosites in RBD mutated from N to Q and 2 O-glycosites mutated from S/T to A, and glycosite 122, 165 and 234 mutated from N to Q).

[0029] In one embodiment, the immunization of the exemplary coronavirus vaccine of the present disclosure, as described herein, results in the accumulation of misfolded S protein in the endoplasmic reticulum. In one embodiment, the immunization of the exemplary coronavirus vaccine of the present disclosure, as described herein, causes the upregulation of BiP/GRP78, XBP1 and p-eIF2 α to induce cell apoptosis and CD8⁺ T-cell response. In one embodiment, the immunization of the coronavirus vaccine of the present disclosure, as described herein, can increase class I major histocompatibility complex (MHC I) expression.

[0030] In some embodiments, the exemplary CoVs described herein includes, but are not limited to, SARS-CoV, MERS-CoV and SARS-CoV-2. In some embodiments, examples of the coronavirus (CoV) described herein include, but are not limited to, alpha-SARS-CoV2, beta-SARS-CoV2, gamma-SARS-CoV2, delta-SARS-CoV2, and omicron-SARS-CoV2 and variants thereof.

[0031] In some embodiments, the present disclosure provides a linear DNA comprising a promoter, 5' untranslated region, 3' untranslated region, expression plasmid with or without S-2P, and poly(A) tail signal sequence, wherein the putative sequon N-X-S/T is changed to Q-X-S/T and the O-glycosite was changed from S/T to A on the expression plasmid.

[0032] In some embodiments, the S-2P expression plasmid comprises the S gene of SARS-CoV-2 encoding the pre-fusion state of the S having proline substitutions of K968 and V969.

[0033] In some embodiments, the present disclosure provides an mRNA, prepared by *in vitro* translation from the above-mentioned DNA.

[0034] In another aspect, the present disclosure provides a vector comprising the modified nucleic acid molecule described above.

[0035] In another aspect, the present disclosure provides a host cell comprising the modified nucleic acid molecule described above.

[0036] In another aspect, the present disclosure provides a modified spike protein described above.

[0037] In another aspect, the present disclosure provides a method for delivery of mRNA for *in vivo* production of a protein comprising: administering to a subject a composition comprising an mRNA of the invention that encodes the protein, wherein the mRNA is encapsulated within a lipid

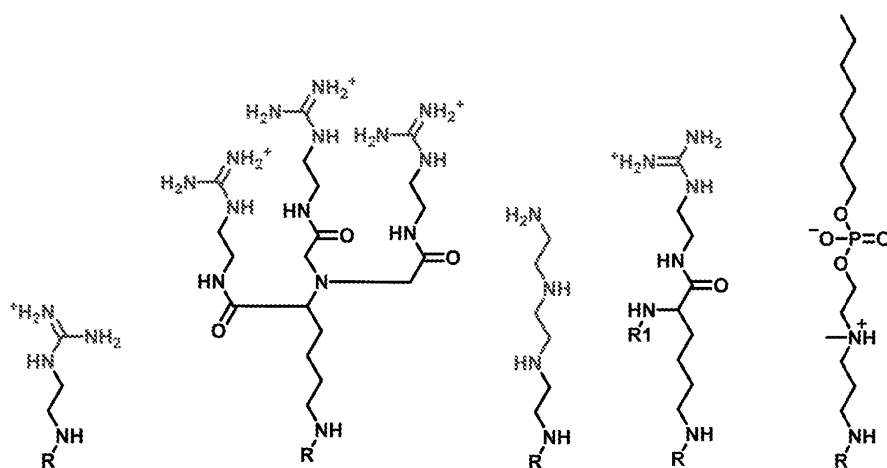
nanoparticle, and wherein the administering of the composition results in the expression of the protein encoded by the mRNA.

[0038] In some embodiments, the mRNA as described herein may be used as the vaccine, either alone or in combination with other vaccines. Accordingly, the present disclosure provides a combo vaccine, comprising the mRNA vaccine of the present disclosure and one or more additional vaccines. The additional vaccine is selected from one or more COVID-19 vaccine, influenza (flu) vaccine, adenovirus vaccine, anthrax vaccine, cholera vaccine, diphtheria vaccine, hepatitis A or B vaccine, HPV vaccine, measles vaccine, mumps vaccine, smallpox vaccine, rotavirus vaccine, tuberculosis vaccine, pneumococcal vaccine and Haemophilus influenzae type b vaccine and any combination thereof.

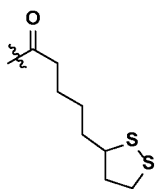
[0039] In another aspect, the present disclosure provides a guanidine-based nanoparticle used as carrier for delivering the modified nucleic acid molecule of any one of claims 1 to 25 to a subject. In one embodiment, the nanoparticle is a liposome or a polymersome.

[0040] In another aspect, the present disclosure provides an mRNA nanocluster comprising the mRNA vaccine as described herein formulated in lipid nanoparticles. In one embodiment, the lipid nanoparticle is a biodegradable lipid nanoparticle.

[0041] In some embodiments, the lipid nanoparticle as described herein is guanidine-based polymers. In one embodiment, the present disclosure provides an mRNA nanocluster, comprising lipid nanoparticles encapsulated with the mRNA vaccine described herein, wherein the lipid nanoparticle comprises guanidine-based polymer units, wherein the guanidine-based as well as zwitterionic groups of the polymer attach to a lipid tail of the polymer, and wherein the guanidine-based polymers adhere to mRNA, thereby forming salt bridges between the guanidinium groups and the phosphates in the mRNA. Examples of guanidine-based polymers include, but are not limited to P1, P2, P3, Pb and Pz as shown below.

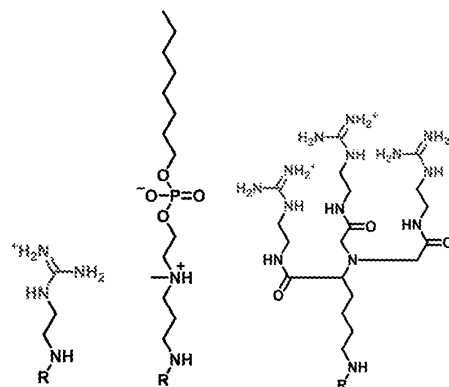
**P1****P2****P3****Pb****Pz**

, wherein R is

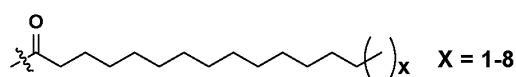


[0042] In some embodiments, the guanidine-based polymer forms a copolymer such as P1/P3 copolymer, P2/P3 copolymer, P1/Pb copolymer, P2/Pb copolymer, P1/Pz copolymer and P2/Pz copolymer.

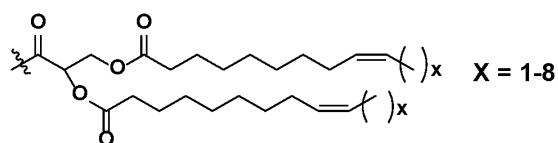
[0043] In some embodiments, the guanidine-based and zwitterionic lipid nanoparticles comprise

**P1****Pz****P2**

, wherein R is



or



[0044] In some embodiments, the mRNA nanocluster described herein has a nanoparticle/mRNA (N/P) ratio of about 1, about 2, about 3, about 4, about 5, about 6, about 7, about 8, about 9, about 10, about 15, about 20, about 30, about 40, about 50, or about 100

[0045] In some embodiments, the mRNA nanocluster described herein has a nanoparticle/mRNA (N/P) ratio of about 10 or about 20.

[0046] In some embodiments, the present disclosure is directed to a nanoparticle/nanocluster composition, comprising a nanoparticle attached with the coronavirus vaccine of the present disclosure. In one embodiment, the nanoparticle is a lipid nanoparticle, a polymeric nanoparticle, an inorganic nanoparticle such as a gold nanoparticle, a liposome, an immune stimulating complex, a virus-like particle, or a self-assembling protein. In a further embodiment, the nanoparticle is a lipid nanoparticle (LNP).

[0047] In some embodiments, the present disclosure provides a vaccine composition comprising the mRNA vaccine, mRNA nanocluster/nanocluster or nanoparticle composition as described herein.

[0048] In some embodiments, the present disclosure is directed to the antibodies and CD8⁺ T cells elicited by the vaccine described herein, which have broader protection against the alpha, beta, gamma, delta and omicron variants.

[0049] In some embodiments, the present disclosure provides a method of immunizing a subject comprising administering the vaccine composition described herein. The present disclosure also provides a method of preventing or treating a coronavirus infection, comprising administering an effective amount of the mRNA vaccine, mRNA nanocluster or mRNA nanoparticle or vaccine composition as described herein to a subject infected with, or at risk of being infected with, a coronavirus. In one embodiment, the mRNA vaccine, mRNA nanocluster or mRNA nanoparticle or vaccine composition as described herein can be used in a method of boosting an adaptive immune response.

[0050] In some embodiments, the mRNA vaccine, mRNA nanocluster or mRNA nanoparticle or vaccine composition as described herein is administered in an initial dose and two, three or four

booster doses. In some embodiments, the mRNA vaccine, mRNA nanocluster or mRNA nanoparticle or vaccine composition is administered in an initial dose and in at least one booster dose about one month, about two months, about three months, about four months, about five months, or about six months following the initial dose. In some embodiments, a provided composition is administered in a second booster dose about six months, about seven months, about eight months, about nine months, about ten months, about eleven months, or about one year following the initial dose.

[0051] In some embodiments, the mRNA vaccine, mRNA nanocluster or mRNA nanoparticle or vaccine composition are administered in one, or more doses. In one embodiment, the dose may include or exclude 5 µg to 50 µg of the mRNA. In some embodiments, the dose is about 5 µg, 10 µg, 15 µg, 20 µg, 25 µg, 30 µg, 35 µg, 40 µg, 45 µg or 50 µg.

[0052] In some embodiments, the mRNA vaccine, mRNA nanocluster or mRNA nanoparticle or vaccine composition are administered via intravenous route, intramuscular route, intradermal route, or subcutaneous route, or by infusion or nasal spray.

[0053] In some embodiments, the present disclosure provides a method for preparing broadly protective vaccines and antibodies against SARS-CoV-2. In one embodiment, the method comprising generating the vaccine using the RNA or DNA of native or glycoengineered S protein whereas the protein expressed within the antigen presentation cells, including the folded or unfolded forms, are processed and presented to T cells.

[0054] These and other aspects will become apparent from the following description of the preferred embodiment taken in conjunction with the following drawings, although variations and modifications therein may be affected without departing from the spirit and scope of the novel concepts of the disclosure.

BRIEF DESCRIPTION OF THE DRAWINGS

[0055] The following drawings form part of the present specification and are included to further demonstrate certain aspects of the present disclosure, the inventions of which can be better understood by reference to one or more of these drawings in combination with the detailed description of specific embodiments presented herein.

[0056] Fig. 1. Conserved epitopes of S protein variants, 10 of which are shielded by Glycans.

[0057] Fig. 2 Identification of N- and O-glycosites and mutations in variants. All 24 glycosites are highly conserved among 6 million S protein sequences.

[0058] Fig. 3. Analysis of S protein expression after transfection with mRNA at 48 hr by western blot. The filter was probed with anti-S and anti- β -actin monoclonal antibodies.

[0059] Fig. 4A to 4F. Humoral immune response in BALB/c mice was shown as serum of anti-S WT (A), S2 (B), RBD (C), deglycosylated S (D), deglycosylated S2 (E) and deglycosylated RBD (F) protein-specific IgG endpoint titer analyzed by ELISA. Mean $\pm$ SD for five independent experiments. *P < 0.001.

[0060] Figs. 5A to 5D. Glycosylation regulated the specificity of elicited antibodies and affected the breadth of mRNA vaccine protection. The alpha (A), beta (B), gamma (C) and delta (D) S protein-specific IgG antibody endpoint titer determined by ELISA. Mean $\pm$ SD for five independent experiments. *P < 0.001, **P < 0.05.

[0061] Figs. 6A to 6F. Neutralization curves of pseudovirus variants are shown in WT (A), alpha (B), beta (C), gamma (D), delta (E), and omicron (F). Mean $\pm$ SD for five independent experiments. *P < 0.001, **P < 0.05.

[0062] Figs. 7A to 7E. Glycosylation affected CD8⁺ T-cell response. Splenocytes isolated from immunized mice were incubated with full-length WT S (A) RBD (B) and S2 (C) peptide pools, then the GrzB-secreting cells were measured by Elispot. The CD4⁺ (D) and CD8⁺ (E) T cells were isolated and incubated with bone-marrow-derived dendritic cells and full-length WT S peptide pool to measure the IFN γ -secreting T cells by flow cytometry. (A-E) Mean $\pm$ SD for five independent experiments. *P < 0.001.

[0063] Figs. 8A to 8F. Glycosylation affects cytokines production. After incubation of the splenocytes isolated from mRNA vaccine immunized mice with full-length WT S peptide pool, (A) IFN γ , (B) IL-2, (C) IL-4, (D) IL-6, (E) IL-12 and (F) IL-13 were measured. (A-F) Mean $\pm$ SD for five independent experiments. *P < 0.001, **P < 0.05.

[0064] Fig. 9. Deletion of glycosites in mRNA to produce deglycosylated S protein and the unfolded protein response. Analysis of deglycosylated S protein expression via HEK293T cells transfected with plasmids and MG132 treatment by western blot. The filter was probed with anti-S and anti-GAPDH monoclonal antibodies.

[0065] Fig. 10. *In vitro* translated deglycosylated S variants in different incubation times as shown in the figure were monitored by ELISA. Mean $\pm$ SD for three independent experiments. *P < 0.001.

[0066] Figs. 11A to 11C. After HEK293 cells were transfected with mRNA vaccine at 48 hrs, the plasma membrane (A), cytosol (without ER) (B) and ER (C) were isolated to analyze the amount

of S protein by western blot. The filter was probed with anti-S, anti-Na/K ATPase, anti-SERCA2 and anti-GAPDH monoclonal antibodies.

[0067] Fig. 12. Analysis of UPR marker proteins BiP/GRP78, XBP1 and p-eIF2 α by western blot after HEK293 cells were transfected with the mRNA vaccine of deglycosylated S protein variants at 48 hrs. The filter was probed with anti-BiP, anti-XBP1, anti-p-eIF2 α and anti- β -actin monoclonal antibodies.

[0068] Fig. 13. To analyze the apoptosis cell via APO-BrdU TUNEL assay after HEK293 cells were transfected with mRNA vaccine at different times as shown in the figure. Mean $\pm$ SD for three independent experiments. *P < 0.001.

[0069] Fig. 14. Analysis of MHC I expression by flow cytometry of DCs after incubation with variants of mRNA vaccines. Mean $\pm$ SD for three independent experiments. *P < 0.001.

[0070] Figs. 15A to 15G. Schematic representation of the SARS-CoV-2 spike and vaccine design: WT (A); S-(deg-RBD) (B); S-(deg-S2) (C); S-(deg-S2-1194) (D); S-(deg-RBD-801) (E); S-(deg-RBD-1194) (F); S-(deg-RBD-122-165-234) (G). NTD, N-terminal domain (14–305 residues). RBD, a receptor-binding domain (319–541 residues). FP, the fusion peptide (788–806 residues). HR1, heptapeptide repeat sequence 1 (912–984 residues). HR2, heptapeptide repeat sequence 2 (1163–1213 residues). TM, transmembrane domain (1213–1237 residues). CT, cytoplasm domain (1237–1273 residues). S2 subunit (686–1273 residues). 2P, (K986P, and V987P). Ψ , the N-glycosylation site; ϕ , O-glycosylation site.

[0071] Fig. 16. Glycosylation on S2 regulated the secretion of soluble pre-fusion SARS-CoV-2 spike protein. After HEK293 cells transfected with the mRNA vaccine that encoded the soluble pre-fusion version of variant S, the location of S was determined by western blot. The filter was probed with anti-S and anti-GAPDH monoclonal antibodies.

[0072] Fig. 17. mRNA vaccine affected MHC II expression on DCs. Analysis of MHC II expression by flow cytometry of DCs after incubation with variants of mRNA vaccines. Mean $\pm$ SD for three independent experiments. *P < 0.001.

[0073] Figs. 18A to 18C. Characterization of the immune response from specific glycosite-deleted S mRNA vaccine. Humoral immune response in BALB/c mice was shown as protein-specific IgG titer from serum against S WT (A), RBD (B) and de-glycosylated RBD (C) analyzed by ELISA.

[0074] Figs. 19A to 19E. Characterization of the immune response from specific glycosite-deleted S mRNA vaccine. Neutralization curves of pseudovirus variants are shown with WT (A), alpha (B), beta (C), gamma (D) and delta (E).

[0075] Figs. 20A to 20B. Characterization of the immune response from specific glycosite-deleted S mRNA vaccine. (A) After incubation of the splenocytes isolated from immunized mice with full-length WT S peptide pools, the GrzB-secreting cells were measured by Elispot. (B) The CD8+ T cells from immunized mice were isolated and incubated with bone-marrow-derived DCs and full-length WT S peptide pool to measure the IFN γ -secreting T cells by flow cytometry. (A-J) Mean $\pm$ SD for five independent experiments. *P < 0.001.

[0076] Figs. 21A and 21B. The protein expression level of specific glycosite-deleted S. (A) Analysis of various S protein expression via HEK293T cells transfected with plasmids and MG132 treatment by western blot. (B) Analysis of S protein expression in HEK293T cells after transfection with mRNA-LNP at 48 hr by western blot. The filter was probed with anti-S and anti-GAPDH monoclonal antibodies.

[0077] Figs. 22A to 22C. Characterization of the immune response from specific glycosite-deleted S mRNA vaccine. After incubation of the splenocytes isolated from immunized mice with RBD (A) and S2 (B) peptide pools, the GrzB-secreting cells were measured by Elispot. (C) The CD4+ T cells were isolated from immunized mice and incubated with bone-marrow-derived DCs and full-length WT S peptide pool to measure the IFN γ -secreting T cells by flow cytometry. (A-C) Mean $\pm$ SD for five independent experiments. *P < 0.001.

[0078] Fig. 23. Propagator **P1** containing guanidine groups and the multivalent display propagator **P2**, facilitate the adherence of mRNA with polymers by forming strong salt bridges between guanidiniums and the phosphates in mRNA.

[0079] Fig. 24. Designed structure of initiators (I0), propagators (P1, P2, Pb, P3, Pz), and the polymer reaction process.

[0080] Figs. 25A and 25B. (A) Agarose gel electrophoresis assay of copolymers with GFP mRNA at N/P ratio = 10. (B) Particle size of mRNA complexes imaged by TEM.

[0081] Figs. 26A to 26C. The fluorescent image of GFP expression of GFP mRNA transfected by poly(disulfide)s in HEK293T cells. (A) GFP mRNA complexed with P1, P3, P1/P3, and PEI in a N/P ratio = 10 (B) GFP mRNA complexed with P1/P3, P2/P3, P1/Pb, and P2/Pb P1/Pz in a N/P ratio = 10 (C) GFP mRNA complexed with different N/P ratio.

[0082] Fig. 27. Agarose gel electrophoresis assay of spike mRNA-polymer complexes at different N/P ratios.

[0083] Fig. 28. Chemiluminescent imaging of spike protein expression mediated by spike mRNA-polymer complexes in HEK293T cells.

[0084] Fig. 29. Cell viability assay of HEK293T cells after treatment of different ratio amounts of polyGu/spike mRNA complexes. The ratio varied from 0.01 to 1. Error bar represent the standard error (mean $\pm$ S.D., n=3).

DETAILED DESCRIPTION OF THE INVENTION

[0085] The terms used in this specification generally have their ordinary meanings in the art, within the context of the invention, and in the specific context where each term is used. Certain terms that are used to describe the invention are discussed below, or elsewhere in the specification, to provide additional guidance to the practitioner regarding the description of the invention. For convenience, certain terms may be highlighted, for example using italics and/or quotation marks. The use of highlighting has no influence on the scope and meaning of a term; the scope and meaning of a term is the same, in the same context, whether or not it is highlighted. It will be appreciated that same thing can be said in more than one way. Consequently, alternative language and synonyms may be used for any one or more of the terms discussed herein, nor is any special significance to be placed upon whether or not a term is elaborated or discussed herein. Synonyms for certain terms are provided. A recital of one or more synonyms does not exclude the use of other synonyms. The use of examples anywhere in this specification including examples of any terms discussed herein is illustrative only, and in no way limits the scope and meaning of the invention or of any exemplified term. Likewise, the invention is not limited to various embodiments given in this specification.

[0086] It must be noted that, as used in this specification and the appended claims, the singular forms "a," "an," and "the" include plural referents unless the content clearly dictates otherwise.

[0087] As used herein, the terms "spike protein" and "spike glycoprotein" and "coronavirus spike protein" are used interchangeable.

[0088] As used herein, the terms "wild-type (native) coronavirus spike protein", "wild-type (native) coronavirus spike glycoprotein", "wild-type (native) spike glycoprotein" and "wild-type (native) spike protein" are used interchangeable.

[0089] As used herein, the terms "treat," "treatment," and "treating" refer to an approach for obtaining beneficial or desired results, for example, clinical results. For the purposes of this disclosure, beneficial or desired results may include inhibiting or suppressing the initiation or progression of an infection or a disease; ameliorating, or reducing the development of, symptoms of an infection or disease; or a combination thereof.

[0090] As used herein, the terms "preventing" and "prevention" are used interchangeably with "prophylaxis" and can mean complete prevention of an infection, or prevention of the

development of symptoms of that infection; a delay in the onset of an infection or its symptoms; or a decrease in the severity of a subsequently developed infection or its symptoms.

[0091] As used herein an "effective amount" refers to an amount of an immunogen sufficient to induce an immune response that reduces at least one symptom of pathogen infection. An effective dose or effective amount may be determined e.g., by measuring amounts of neutralizing secretory and/or serum antibodies, e.g., by plaque neutralization, complement fixation, enzyme-linked immunosorbent (ELISA), or microneutralization assay.

[0092] As used herein, the term "vaccine" refers to an immunogenic agent (with or without an adjuvant), such as an immunogen derived from a coronavirus, which is used to induce an immune response against the coronavirus that provides protective immunity (e.g., immunity that protects a subject against infection with the coronavirus and/or reduces the severity of the condition caused by infection with the coronavirus). The protective immune response may include formation of antibodies and/or a cell-mediated response. Depending on context, the term "vaccine" may also refer to a suspension or solution of an immunogen that is administered to a subject to produce protective immunity.

[0093] As used herein, the term "subject" includes humans and other animals. Typically, the subject is a human. For example, the subject may be an adult, a teenager, a child (2 years to 14 years of age), an infant (birth to 2 year), or a neonate (up to 2 months). In particular aspects, the subject is up to 4 months old, or up to 6 months old. In some aspects, the adults are seniors about 65 years or older, or about 60 years or older. In some aspects, the subject is a pregnant woman or a woman intending to become pregnant. In other aspects, subject is not a human; for example, a non-human primate; for example, a baboon, a chimpanzee, a gorilla, or a macaque. In certain aspects, the subject may be a pet, such as a dog or cat.

[0094] As used herein, the term "pharmaceutically acceptable" means being approved by a regulatory agency of a U.S. Federal or a state government or listed in the U.S. Pharmacopeia, European Pharmacopeia or other generally recognized pharmacopeia for use in mammals, and more particularly in humans. These compositions can be useful as a vaccine and/or antigenic compositions for inducing a protective immune response in a vertebrate.

[0095] The outbreak of SARS-CoV-2 that causes COVID-19 has resulted in a global pandemic. The current clinical management for SARS-CoV-2 infection includes prevention, control measures, and supporting care. To contain the current pandemic and possible future reoccurrence, it is important to better understand this virus and to develop rapid diagnosis methods, therapeutic treatments, and preventive vaccines to combat such dangerous pathogens. Most of the

vaccine and antibody development efforts are mainly focused on the extensively glycosylated SARS-CoV-2 S protein, the important mediator of virus entry to the host cell by binding to the angiotensin converting enzyme 2 (ACE2) receptor on host cell surface. Like many other viral fusion proteins, the SARS-CoV-2 S protein utilizes a glycan coat to shield the S protein backbone in both pre-fusion and post-fusion conformation and evade the host immune response. However, how the post-translational modification affects the translated immunogen after mRNA vaccination is still unknown, and among the posttranslational modification events, glycosylation plays an important role in the regulation of protein folding, structure and function. This present disclosure is aimed at developing mono-GlcNAc decorated and glycosite-engineered variants (removal of a non-essential glycosite via reverse genetics to replace Asn with Gln) for full length S protein and its subunits including S1 or S2, and the RBD domain as vaccine candidates for immunization studies to generate antigen-specific neutralizing antibodies.

[0096] It is believed that the development of innovative strategies and broadly protective vaccines to combat CoV infections can lead to important discoveries with medical significance but was not emphasized otherwise. The principles and strategies developed in the present disclosure provide universal coronavirus mRNA vaccines against different CoVs and their variants.

[0097] Immunogenic peptide Derived from Spike Protein of Coronavirus

[0098] To date, more than 8-million sequences of S protein with over 1,000 sites of mutation in its 1,273 amino-acid sequence have been reported, including the highly contagious D614G mutant and that from the UK and South Africa. In addition, all the glycosites on S protein are highly conserved and the conserved peptide epitopes on S protein are largely shielded by glycans. This poses a major challenge in the development of broadly effective antibodies and vaccines to combat upcoming viral strains. The present disclosure develops a more effective vaccine design strategy using the S protein with engineered glycosylation as immunogens to better expose the highly conserved epitopes for vaccine design in order to elicit broadly protective immune responses.

[0099] The present disclosure found that removal of glycan shields on viral surface glycoproteins to expose more conserved epitopes is a very effective approach for vaccine design against SARS-CoV-2. Because the single GlcNAc residue linked to Asn is the minimum component of the *N*-glycan required for glycoprotein folding and stabilization, it is therefore postulated that trimming of *N*-glycans to leave a single GlcNAc on SARS-CoV-2 S protein will not affect its folding but will facilitate the maximum exposure of protein backbone to elicit robust and protein specific immune response while maintaining its structural integrity.

[00100] By removing glycan shields on the spike protein of SARS-CoV-2, the present disclosure provides an immunogenic peptide, comprising at least one amino acid sequence selected from a group consisting of: TESIVRFPNITNL (SEQ ID NO: 41), NITNLCPFGVEVFNATR (SEQ ID NO: 42), LYNSASFSTFK (SEQ ID NO: 43), LDSKVGGNYN (SEQ ID NO: 44), KSNLKPFERDIST (SEQ ID NO: 45), KPFERDISTEIYQAG (SEQ ID NO: 46), GPKKSTNLVKKNC (SEQ ID NO: 47), NCDVVIGIV[N]NTVY (SEQ ID NO: 48), PELDSFKEELDKYFK[N]HTS (SEQ ID NO: 49), VNIQKEIDRLNEVA (SEQ ID NO: 50), NL[N]ESLIDLQ (SEQ ID NO: 51) and LGKYEQYIKWP (SEQ ID NO: 52) or an amino acid sequence having at least about 99%, 98%, 97%, 96%, 95% or 90% identity to any of SEQ ID NOs: 41 to 52.

[00101] In some embodiments, the immunogenic peptide comprises at least one amino acid sequence selected from a group consisting of SEQ ID NOs: 41 to 43 and 45 to 51.

[00102] The amino acid sequence of SEQ ID NOs: 41 to 52, individual or in combination, can be used as antigen(s) capable of stimulating an immune response against coronaviruses.

[00103] Conventional methods, e.g., chemical synthesis or recombinant technology, can be used to make the immunogenic peptide as described herein.

[00104] The immunogenic peptide or an expression vector capable of expressing the immunogenic peptide can be mixed with a pharmaceutically acceptable carrier to form an immunogenic composition. The composition can be administered to a subject in need thereof to prevent or treat coronavirus infection.

[00105] The composition can be formulated with a pharmaceutically acceptable carrier such as a phosphate buffered saline, a bicarbonate solution, and/or an adjuvant. Suitable pharmaceutical carriers and diluents, as well as pharmaceutical necessities for their use, are known in the art. This composition may be prepared as an injectable, liquid solution, emulsion, or another suitable formulation.

[00106] Examples of adjuvants include, but are not limited to, alum-precipitate, Freund's complete adjuvant, Freund's incomplete adjuvant, CpG, QS21, monophosphoryl-lipid A/trehalose dicorynomycolate adjuvant, water in oil emulsion containing *Corynebacterium parvum* and tRNA, and other substances that accomplish the task of increasing immune response by mimicking specific sets of evolutionarily conserved molecules including liposomes, lipopolysaccharide (LPS), molecular cages for antigen, components of bacterial cell walls, and endocytosed nucleic acids such as double-stranded RNA, single-stranded DNA, and unmethylated CpG dinucleotide-containing DNA. Other examples include cholera toxin, *E. coli* heat-labile enterotoxin, liposome,

immune-stimulating complex (ISCOM), immunostimulatory sequences oligodeoxynucleotide, and aluminum hydroxide. The composition can also include a polymer that facilitates in vivo delivery.

[00107] Coronavirus mRNA Vaccines

[00108] Coronaviruses (CoVs) infect human and animals and cause varieties of diseases, including respiratory, enteric, renal, and neurological diseases. CoV uses its spike glycoprotein (S), a main target for neutralization antibody, to bind its receptor, and mediate membrane fusion and virus entry. The coronavirus spike protein is highly conserved among all human coronaviruses (CoVs) and is involved in receptor recognition, viral attachment, and entry into host cells. Similarly, SARS-CoV-2 S protein is also highly conserved with that of CoVs. The SARS-CoV-2 S protein has three major immunogenic domains: the N-terminal domain (NTD), the receptor binding domain (RBD) and the subunit 2 domain (S2). Previous studies have shown that the neutralizing antibodies (NAbs) that recognize the RBD are highly protective against SARS-CoV-2 and other coronaviruses, and the S protein is highly glycosylated (24 glycosites per monomer) and frequently mutated with millions of sequences reported by GISAID. The most conserved regions of SARS-CoV-2 S protein are located in the RBD and S2 domains, which are largely shielded by glycans (*Han-Yi Huang, et al. Impact of glycosylation on a broad-spectrum vaccine against SARS-CoV-2. bioRxiv preprint. doi: www.biorxiv.org/content/10.1101/2021.05.25.445523v2.full*), and antibodies recognized these regions could provide a broad protection against variants of SARS-CoV-2 (*Maximilian M Sauer, et al. Structural basis for broad coronavirus neutralization. Nat Struct Mol Biol. 28 (6):478-486 (2021)1314*; C.; Wang. et al. A conserved immunogenic and vulnerable site on the coronavirus spike protein delineated by cross-reactive monoclonal antibodies. *Nat Commun. 12 (1):1715 (2021)*). Glycosylation on the target antigens or pathogens regulated the induction of antibody, but whether glycosylation affected T cell response was still unknown. In fact, it is difficult or impossible to express the S protein from the plasmid with deletion of certain glycosylation sites (*Han-Yi Huang, et al. Impact of glycosylation on a broad-spectrum vaccine against SARS-CoV-2. bioRxiv preprint. doi: www.biorxiv.org/content/10.1101/2021.05.25.445523v2.full*).

[00109] The present disclosure surprisingly found that using the mRNA technology to remove the glycan shields to better expose the conserved regions is an effective strategy of broad-spectrum vaccine design. In the present disclosure, the mRNA of coronavirus spike protein (such as SARS-CoV-2 S protein) with mutation of specific glycosites is used as a model for immunization in order to investigate how the glycosite-mutated mRNA affects the protein expression and immune response.

[00110] Accordingly, the present disclosure provides a modified nucleic acid molecule encoding a modified spike protein comprising one or more amino acid substitutions of asparagine (N) to glutamine (Q) at N-linked glycosylation sequons (N-X-S/T), wherein X is any amino acid residue except proline, and S/T denotes a serine or threonine residue.

[00111] The modified nucleic acid molecule can be an mRNA or a single or double-strand DNA and used as immunogen or vaccine against a pathogen. In one embodiment, the pathogen is CoV. Examples of the CoV include, but are not limited to, SARS-CoV, MERS-CoV and SARS-CoV-2. Examples of the SARS-CoV-2 include, but are not limited to, alpha-SARS-CoV2, beta-SARS-CoV2, gamma-SARS-CoV2, delta-SARS-CoV2, and omicron-SARS-CoV2 and variants thereof.

[00112] Compared with the wild type spike protein of Wuhan strains and delta strains (such as SEQ ID NOs: 2, 16, 18 and 20), the modified spike protein described herein comprises one or more amino acid deletions or additions at N-linked glycosylation sequons (N-X-S/T) to eliminate N-linked glycan sequons. Alternatively, the modified spike protein described herein comprises one or more amino acid substitutions of S/T to alanine (A) at O-linked glycosylation sites to eliminate O-linked glycosylation sites.

[00113] The mRNA of a coronavirus spike protein can be used as a coronavirus vaccine, which has mutation of one or more glycosites in the receptor binding domain (RBD), the subunit 1 (S1) or the subunit 2 (S2) domain, or a variant thereof.

[00114] The mutation of CoV or a variant thereof as described herein can be deletion, addition or substitution. In some embodiments, a coronavirus spike protein mRNA has one or more mutation of the glycosites in RBD, S1 or S2 with one or more replacements of N to Q or S/T to A, or a combination thereof. The mutation of the N-glycosites is to change the putative sequon N-X-S/T to Q-X-S/T and/or change S/T of the O-glycosite to A.

[00115] The glycosites with N to Q replacement include, but are not limited to, the following:

an S-(deg-RBD) which is S protein with all 2 N-glycosites in RBD mutated from N to Q and 2 O-glycosites mutated from S/T to A (such as SEQ ID NO: 4, 22, 24 or 26);

an S-(deg-S2) which is S protein with all 9 glycosites in S2 mutated from N to Q (such as SEQ ID NO: 6, 28, 30 or 32);

an S-(deg-S2-1194) which is S protein with 8 glycosites in S2 mutated from N to Q, except glycosite 1194 (such as SEQ ID NO: 8 or 34);

an S-(deg-RBD-801) which is S protein with all 2 N-glycosites in RBD mutated from N to Q and 2 O-glycosites mutated from S/T to A, and glycosite 801 mutated from N to Q (such as SEQ ID NO: 10 or 36));

an S-(deg-RBD-1194) which is S protein with all 2 N-glycosites in RBD mutated from N to Q and 2 O-glycosites mutated from S/T to A, and glycosite 1194 mutated from N to Q (such as SEQ ID NO: 12 or 38)); and

an S-(deg-RBD-122-165-234) which is S protein with all 2 N-glycosites in RBD mutated from N to Q and 2 O-glycosites mutated from S/T to A, and glycosite 122, 165 and 234 mutated from N to Q (such as SEQ ID NO: 14 or 40).

[00116] In a further embodiment, the mRNA or DNA for S-(deg-RBD) has the sequence of SEQ ID NO: 3, 21, 23 or 25, the mRNA or DNA for S-(deg-S2) has the sequence of SEQ ID NO: 5, 27, 29 or 31, the mRNA or DNA for S-(deg-S2-1194) has the sequence of SEQ ID NO: 7 or 33, the mRNA or DNA for S-(deg-RBD-801) has the sequence of SEQ ID NO: 9 or 35, the mRNA or DNA for S-(deg-RBD-1194) has the sequence of SEQ ID NO: 11 or 37, the mRNA or DNA for S-(deg-RBD-122-165-234) has the sequence of SEQ ID NO: 13 or 39.

[00117] The present disclosure provides a linear DNA comprising a promoter, 5' untranslated region, 3' untranslated region, expression plasmid with or without S-2P, and poly(A) tail signal sequence, wherein the putative sequon N-X-S/T is changed to Q-X-S/T and the O-glycosite was changed from S/T to A on the expression plasmid. In one embodiment, the S-2P expression plasmid comprises the S gene of SARS-CoV-2 encoding the pre-fusion state of the S having proline substitutions of K968 and V969.

[00118] The mRNA can be prepared by *in vitro* translation from the above-mentioned DNA using a vector comprising the modified nucleic acid molecule and a host cell comprising the vector as described herein. The target spike protein gene is synthetically manufactured and inserted into in a plasmid, or a small, circular piece of DNA. Plasmids are used in mRNA vaccine production because they are easy to replicate (copy) and reliably contain the target gene sequence. The two strands of plasmid DNA are separated. Then, RNA polymerase, the molecule that transcribes RNA from DNA, uses the spike protein gene to create a single mRNA molecule. Finally, other molecules break down the rest of the plasmid to ensure that only the mRNA is packaged as a vaccine. The speed and efficiency of this process can make large amounts of mRNA in a short period of time.

[00119] The present disclosure has found that immunization of wild-type S protein with the glycans at all N-glycosites trimmed down to N-acetylglucosamine (GlcNAc) as the mono-GlcNAc

decorated S protein (S_{mg}) induced broadly protective antibody and $CD4^+$ as well as $CD8^+$ T cell responses against the variants of concern, including the alpha, beta, gamma, delta, and omicron variants. Further study shows that most of the conserved epitopes on S protein are located in the RBD and the HR2 domain of the S2 subunit, but these conserved epitopes are largely shielded by glycans to escape the immune response. So, removal of the shielded glycans will expose more conserved epitopes and induce broader and stronger immune responses. The present disclosure also uses the single B cell technology to screen the B cells from S_{mg} immunized mice to identify a broadly neutralizing monoclonal antibody that targets the highly conserved region in RBD which was not induced in the immunization of fully glycosylated S protein, further demonstrating that removal of glycan shields from S protein is an effective strategy for development of broadly protective vaccine against SARS-CoV-2 variants. To translate this finding into the mRNA vaccine design, here we focus on the study of SARS-CoV-2 spike mRNA with mutation of specific glycosites in RBD, S1 or S2 or a combination thereof with N to Q and S/T to A replacement and investigation of their protein expression and immune response as well as breadth of protection.

[00120] Immunization of such mRNA results in the accumulation of misfolded spike protein in the endoplasmic reticulum and causes the upregulation of BiP/GRP78, XBP1 and p-eIF2 α to induce cell apoptosis and CD8 T-cell response. In addition, dendritic cells (DCs) incubated with S2 glycosite-deleted mRNA vaccine increased class I major histocompatibility complex (MHC I) expression. Furthermore, removing the glycosites that affected the stability of spike protein, decreased antibody production and increased CD8 $^+$ T-cell response. The present disclosure provides broad-spectrum mRNA vaccines which would not be achieved using expressed proteins as antigens.

[00121] mRNA Nanocluster and Nanoparticles

[00122] In one aspect, the present disclosure provides an mRNA nanocluster comprising the mRNA vaccine as described herein formulated in a lipid nanoparticle.

[00123] A biodegradable lipid nanoparticle can be used as the lipid nanoparticle. In one embodiment, the biodegradable lipid nanoparticle is guanidine-based polymers.

[00124] In another embodiment, the present disclosure provides an mRNA nanocluster, comprising a biodegradable lipid nanoparticles encapsulated with the mRNA vaccine described herein, wherein the biodegradable lipid nanoparticle comprises guanidine-based and zwitterionic units, wherein the guanidine-based as well as zwitterionic groups attach to a lipid tail of the polymer, and wherein the guanidine-based groups adhere to mRNA, thereby forming salt bridges

between the guanidinium groups and the phosphates in the mRNA. Examples of guanidine-based polymers include, but are not limited to P1, P2, P3, Pb and Pz as described herein.

[00125] The disclosure provides a guanidine-based lipid nanoparticle as carrier for mRNA nanovaccine formulation. The polymers generate an efficient delivery of mRNA to antigen presenting cells, showing a strong ability of endosomal escape. The timely degradation of poly(disulfide)s by intracellular glutathione also minimizes the cytotoxicity as compared to other nondegradable nanocarriers.

[00126] In another embodiment, the mRNA nanocluster has a nanoparticle/mRNA (N/P) ratio of about 10 or about 20.

[00127] The coronavirus mRNA vaccine of the present disclosure can also be attached to a nanoparticle.

[00128] The mRNA nanocluster and nanoparticles are particles between 1 and 100 nanometers (nm) in size which can be used as a substrate for immobilizing ligands. The nanoparticle may, for example, be a lipid nanoparticle, a polymeric nanoparticle, an inorganic nanoparticle such as a gold nanoparticle, a liposome, an immune stimulating complex (ISCOM), a virus-like particle (VLP), or a self-assembling protein.

[00129] In one embodiment, lipid nanoparticle formulations typically comprise one or more lipids. In some embodiments, the lipid is a cationic or an ionizable lipid. In some embodiments, lipid nanoparticle (LNP) formulations further comprise other components, including a phospholipid, a structural lipid, a quaternary amine compound, and a molecule capable of reducing particle aggregation, for example a PEG or PEG-modified lipid.

[00130] The mRNA vaccine as described herein can be encapsulated in liposome or polymersome.

[00131] The conventional liposome is manufactured by the way of forming lipid bimolecular membrane in one step, therefore, it is common that both inside and outside membrane of the liposome are made of the same constituting component. Examples of the ingredient of liposome include, but are not limited to, DSPC, DOTAP, DMG, PEGylated DMG, cholesterol and combination thereof. In one embodiment, mRNA liposome is produced by mixing the mRNA and lipid ingredient at a ratio as described herein at room temperature.

[00132] Polymersomes, as disclosed herein, are enclosures, self-assembled from amphiphilic block copolymers. These amphiphilic block copolymers are macromolecules comprising at least one hydrophobic polymer block and at least one hydrophilic polymer block.

When hydrated, these amphiphilic block copolymers self-assemble into enclosures such that the hydrophobic blocks tend to associate with each other to minimize direct exposure to water and form the inner surface of the enclosure, and the hydrophilic blocks face outward, forming the outer surface of the enclosure. The hydrophobic core of these aqueous soluble polymersomes may provide an environment to solubilize additional hydrophobic molecules. As such, these aqueous soluble polymersomes may act as carrier polymers for hydrophobic molecules encapsulated within the polymersomes. Moreover, the self-assembly of the amphiphilic block polymers occurs in the absence of stabilizers, which would otherwise provide colloidal stability and prevent aggregation. In one embodiment, mRNA liposome is produced by mixing the mRNA and the polymer at a ratio as described herein at room temperature.

[00133] Vaccine, Combo Vaccine, Vaccine Compositions, Methods and Therapeutic Use

[00134] The mRNA as described herein may be used as the vaccine, either alone or in combination with other vaccines. Accordingly, the present disclosure provides a combo vaccine, comprising the mRNA vaccine of the present disclosure and one or more additional vaccines. The additional vaccine is selected from one or more COVID-19 vaccine, influenza (flu) vaccine, adenovirus vaccine, anthrax vaccine, cholera vaccine, diphtheria vaccine, hepatitis A or B vaccine, HPV vaccine, measles vaccine, mumps vaccine, smallpox vaccine, rotavirus vaccine, tuberculosis vaccine, pneumococcal vaccine and Haemophilus influenzae type b vaccine and any combination thereof.

[00135] The present disclosure also provides a vaccine composition comprising an mRNA vaccine, mRNA nanocluster or mRNA nanoparticle as described herein. The present disclosure also provides a method of preventing or treating a coronavirus infection, comprising administering an mRNA vaccine, mRNA nanocluster or mRNA nanoparticle or a vaccine composition as described herein to a subject. In one embodiment, the subject is infected with, or at risk of being infected with, a coronavirus.

[00136] The mRNA vaccine, mRNA nanocluster or mRNA nanoparticle or vaccine composition as described herein can be administered in an initial dose and two, three or four booster doses. In some embodiments, the mRNA vaccine, mRNA nanocluster or mRNA nanoparticle or vaccine composition is administered in an initial dose and in at least one booster dose about one month, about two months, about three months, about four months, about five months, or about six months following the initial dose. In some embodiments, a provided composition is administered in a second booster dose about six months, about seven months, about

eight months, about nine months, about ten months, about eleven months, or about one year following the initial dose.

[00137] The mRNA vaccine, mRNA nanocluster or mRNA nanoparticle or vaccine composition are administered in one, or more doses. In one embodiment, the dose may include or exclude 5 µg to 50 µg. In some embodiments, the dose is about 5 µg, 10 µg, 15 µg, 20 µg, 25 µg, 30 µg, 35 µg, 40 µg, 45 µg or 50 µg.

[00138] The vaccine composition preferably comprises a pharmaceutically acceptable vaccine, carrier or diluent. The vaccine composition may be formulated using any suitable method. Formulation of with standard pharmaceutically acceptable carriers and/or excipients may be carried out using routine methods in the pharmaceutical art. The exact nature of a formulation will depend upon several factors including the vaccine to be administered and the desired route of administration. Suitable types of formulation are fully described in Remington's Pharmaceutical Sciences, 19th Edition, Mack Publishing Company, Eastern Pennsylvania, USA.

[00139] The vaccine composition or pharmaceutical composition as described herein may be administered by any route. Suitable routes include, but are not limited to, the nasal, intravenous, intramuscular, intraperitoneal, subcutaneous, intradermal, transdermal and oral/buccal routes.

[00140] Compositions may be prepared together with a physiologically acceptable carrier or diluent. Typically, such compositions are prepared as liquid suspensions of nanoparticles. The nanoparticles may be mixed with an excipient which is pharmaceutically acceptable and compatible with the active ingredient. Suitable excipients are, for example, water, saline, dextrose, glycerol, of the like and combinations thereof.

[00141] Suitable compositions wherein the carrier is a liquid, for administration as for example a nasal spray or as nasal drops, include aqueous or oily solutions of the active ingredient. Formulations suitable for aerosol administration may be prepared according to conventional methods and may be delivered with other therapeutic agents.

[00142] In addition, if desired, the pharmaceutical compositions may contain minor amounts of auxiliary substances such as wetting or emulsifying agents, and/or pH buffering agents.

[00143] Without intent to limit the scope of the invention, exemplary instruments, apparatus, methods and their related results according to the embodiments of the present invention are given below. Note that titles or subtitles may be used in the examples for convenience of a reader, which in no way should limit the scope of the invention. Moreover, certain theories are proposed and disclosed herein; however, in no way they, whether they are right or wrong, should

limit the scope of the invention so long as the invention is practiced according to the invention without regard for any particular theory or scheme of action.

Examples

[00144] Materials and Methods

[00145] **Cell lines.** The Human embryonic kidney cells (HEK293) were maintained in Dulbecco's modified Eagle's medium (DMEM) (Invitrogen, Rockville, MD) with 10% heat-inactivated fetal bovine serum (FBS) (Thermo Scientific) and antibiotics (100 U/ml penicillin G and 100 gm/ml streptomycin).

[00146] **Antibodies and proteins.** The rabbit anti-SARS-CoV-2 S polyclonal antibody, and SARS-CoV-2 full length S, S2, RBD and variant proteins (293T cell expressed) were purchased from Sino Biologicals (Beijing, China). Mouse monoclonal anti- β -actin, GAPDH and rabbit monoclonal anti-MHCII antibodies were purchased from Millipore. The rabbit monoclonal anti-Na/K ATPase was obtained from ABcan. The mouse monoclonal anti-SERCA2 and rabbit monoclonal anti-MHC I antibodies was obtained from Invitrogen. The rabbit monoclonal anti-BiP/GRP78, XBP1 and p-eIF2 α antibodies were purchased from ABclonal. All commercial antibodies were validated for specificity by companies and us via western blot. To obtain the deglycosylated protein, S, RBD, S1 or S2 protein was deglycosylated in a buffer solution with PNGase F (Sigma) at 37°C for 24 h in the dark. After deglycosylation, samples were purified and checked by Western blot.

[00147] **mRNA vaccine of deglycosylated S protein and formulation.** The pre-fusion state of the S, the codon-optimized S gene of SARS-CoV-2 was synthesized by GenScript and cloned into pcDNA3.1 or pVax, and in one embodiment was stabilized by proline substitutions of K968 and V969 (S-2P). The soluble version of S ended with glutamine Q1208 of S-2P followed by a T4 fibrin (foldon) trimerization motif, thrombin cleavage site and 6xHis tag at the C-terminus was constructed. To mutate the N-glycosites, the putative sequon N-X-S/T was changed to Q-X-S/T and the O-glycosite was changed from S/T to A by using site-directed mutagenesis on the S-2P expression plasmid. To obtain the mRNA vaccine, the linear DNA that contained the T7 promoter, 5' untranslated region, 3' untranslated region, S-2P, and poly(A) tail signal sequence was amplified by using TOOLS Ultra High Fidelity DNA Polymerase (BIOTOOLS Co., Ltd., Taipei, Taiwan) with 1 μ l of the DNA template in an mMESSAGE mMACHINE® Kit (Thermo Scientific) at 37 °C for 1hr according to the manufacturer's protocol. The mRNA was purified by RNA cleanup kit (BioLabs), according to the manufacturer's protocol and stored at -80 °C until further use. For the formulation mRNA-LNP, mRNA was encapsulated in LNP using a self-

assembly process in which an aqueous solution of mRNA at pH 4.0 was rapidly mixed with an ethanolic lipid mixture containing ionizable cationic lipid, phosphatidylcholine, cholesterol and polyethylene glycol-lipid. The compositions of LNP were DSPC (Sigma), cholesterol (Sigma), DOTAP (Sigma) and DMG-PEG 2000 (Sigma). The mRNA-LNP was characterized and subsequently stored at -80°C at a concentration of 1 mg/ml. After HEK293 cells were transfected with 10 µg of mRNA-LNP in six wells of a plate at 48 hrs, the total cell lysate was collected to monitor the expression of S by western blot.

[00148] Animals and immunizations. BALB/c mice aged 6–8 weeks old (n = 5) were immunized intramuscularly with 50 µg mRNA-LNP in PBS with 300 mM sucrose. Animals were immunized at week 0, boosted with a second vaccination at week 2, and serum samples and spleens were collected from each mouse in one week after the booster immunization. The animal experiments were evaluated and approved by the Institutional Animal Care and Use Committee of Academia Sinica.

[00149] Serum IgG titer measure. Anti-S protein ELISA was used to determine IgG titer. Plates were coated with 50 ng/well of variant S protein as shown in Fig. 4 and 5, and then blocked with 5% skim milk. The serum from immunized mice and HRP-conjugated secondary antibody were sequentially added. Peroxidase substrate solution (TMB) and 1M H₂SO₄ stop solution were used and absorbance (OD 450 nm) was read by a microplate reader.

[00150] Pseudovirus neutralization assay for serum study. Pseudovirus was constructed by the RNAi Core Facility at Academia Sinica. Briefly, the pseudotyped lentivirus carrying SARS-CoV-2 S protein or variant was generated by transiently transfecting HEK-293T cells with pCMV-ΔR8.91, pLAS2w.Fluc. Ppuro and pcDNA3.1-nCoV-SA18. HEK-293T cells were seeded one day before transfection followed by delivery of plasmids into cells by TransITR-LT1 transfection reagent (Mirus). The culture medium was refreshed at 16 h and harvested at 48 h and 72h post-transfection. Cell debris was removed by centrifugation and the supernatant was passed through 0.45 µm syringe filter (Pall Corporation). The pseudotyped lentivirus was then stored at -80°C. To estimate the lentiviral titer by AlarmaBlue assay (Thermo Scientific), the transduction unit (TU) of pseudotyped lentivirus was estimated by using cell viability assay. HEK-293T cells expressing human ACE2 gene were plated on 96-well plate one day before lentivirus transduction. To determine the titer of pseudotyped lentivirus, different amounts of lentivirus were added into the culture medium containing polybrene (final concentration 8 µg/ml) (sigma) and spin infection was carried out at 1,100 xg in 96-well plate for 30 min at 37°C. After incubation for 16 h, the culture medium containing virus and polybrene was removed and replaced with fresh complete DMEM containing 2.5 µg/ml puromycin (sigma). After treating puromycin for 48 h, the culture

medium was removed and the cell viability was detected by using AlarmaBlue reagents according to manufacturer's instruction. The survival rate of uninfected cells was set as 100%, and the virus titer was determined by plotting the survival cells versus diluted viral dose.

[00151] For neutralization assay, heat-inactivated sera or antibodies were serially diluted with desired dilution and incubated with 1,000 TU of SARS-CoV-2 pseudotyped lentivirus in DMEM for 1 h at 37°C. The mixture was then inoculated with 10,000 HEK-293T cells stably expressing human ACE2 gene in 96-well plate. The culture medium was replaced with fresh complete DMEM (supplemented with 10% FBS and 100 U/ml Penicillin/ Streptomycin) at 16 h post-infection and continuously cultured for another 48 h. The expression level of luciferase gene was determined by using Bright-Glo™ Luciferase Assay System (Promega). The relative light unit (RLU) was detected by Tecan i-control (Infinite 500). The percentage of inhibition was calculated as the ratio of RLU reduction in the presence of diluted serum to the RLU value of no serum control and the calculation formula was shown below: $(RLU_{\text{control}} - RLU_{\text{Serum}}) / RLU_{\text{control}}$.

[00152] Informatic analysis of SARS-CoV-2 S protein. The 1,117,474 S protein sequences of SARS-CoV-2 and their variants were extracted from the Global Initiative on Sharing Avian Influenza Database (GISAID version: Apr. 18, 2021). The S-protein 3D structure model with representative glycan profile was constructed by CHARMM-GUI and OpenMM programs. The transmembrane region of the spike protein defined by UniProt was used in this study. The input of CHARMM-GUI includes the PDB file 6VSB_1_1_1, the representative glycan profile, and parameter settings. Relative solvent accessibility (RSA) of the spike protein with and without representative glycans are calculated by the FreeSASA program. The probe radius 7.2Å was used in the FreeSASA program to mimic the average size of the hypervariable loops in a complementarity determining region (CDR) of an antibody. The RSA value of each residue used in this study was the average RSA value from three protein chains. The definition of exposed/buried residues was the same as the study by Kajander, T. et al.

[00153] Measurement of GrzB and IFN γ secreting cells. A total of 5×10^5 splenocytes from immunized mice were ex vivo restimulated with full-length S, RBD and S2 peptide mix (0.1 μ g/ml final concentration per peptide) (Sino Biologicals) in the GrzB ELISpot assays (R&D Systems) according to the manufacturer's instructions and spots were counted. For T cell subtyping, CD8⁺ T cells and CD4⁺ T cells were isolated from splenocyte suspensions using Dynabeads Untouched Mouse CD4 and CD8 Cells kit (Invitrogen) according to the manufacturer's instructions. CD4⁺ or CD8⁺ T cells (1×10^5) were subsequently restimulated with 5×10^4 syngeneic bone-marrow-derived DCs loaded with full-length WT S peptide mix (0.1 μ g/ml final concentration) (Sino Biologicals). The purity of isolated T cell subsets was determined by

flow cytometry to calculate the spot counts per 1×10^5 CD4⁺ or CD8⁺ T cells. For flow cytometry, cells were suspended in FACS buffer [2% (vol/vol) FBS in PBS] at a density of 10^6 cells/ml and the antibody used in this study was anti-IFN γ (abcam). Cellular fluorescence intensity was analyzed by FACS Canto (BD Biosciences) and FCS Express 3.0 software.

[00154] Measurement of IFN γ and other cytokines. IFN γ , IL-2, IL-4, IL-6, IL-12, and IL-13 were measured by using ELISA kit according to the manufacturer's protocol (IFN- γ : Boster Biological Technology Co., Ltd; IL-2, IL-4, IL-6, IL-12, and IL-13: R&D Systems).

[00155] DNA plasmid transfection and MG132 treatment. After HEK293 cell seeded in the 6 well plate, cells were transfected with 3 μ g of each plasmid by TransIT®-LT1 Transfection Reagent (Mirus) and then incubated with 1 μ M MG-132 (MedChemExpress) or DMSO at 37°C for 24 h. The total lysate was collected and the variant S expression was analyzed by western blot.

[00156] *In vitro* translation. The *in vitro* translation was performed with the plasmid that encoded S-2P using Glycoprotein Expression in a Human IVT System (Thermo) according to the manufacturer's instructions. The expression of S protein in different incubation time periods was monitored by SARS-COV-2 spike protein ELISA kit (ABclonal) according to the manufacturer's protocol.

[00157] Unfolded protein response detection. After HEK293 cells were transfected with 10 μ g of S mRNA with TransIT®-mRNA Transfection Kit (Mirus) for 48 hrs, the plasma membrane and ER were isolated by Minute™ ER Enrichment Kit (Invent Biotech) according to the manufacturer's protocol. The S protein in the plasma membrane, cytosol and ER was analyzed by western blot. Total lysate was collected and the UPR markers XBP1, BiP/GRP78 and p-eIF2 α were monitored by western blot. The apoptosis cells were measured by APO™-BrdU TUNEL Assay Kit (Thermo) according to the manufacturer's instructions.

[00158] The soluble version of S-2P expression. HEK293 cells were transfected with 10 μ g of mRNA that encoded the soluble version of variant S-2P with TransIT®-mRNA Transfection Kit (Mirus) for 72 hrs, the S protein was purified from the cell supernatants using Ni-NTA affinity column (GE Healthcare). The purified protein and total lysate were monitored for the protein level of S by western blot.

[00159] mRNA vaccine induced MHC I / II expression on DCs. DCs were isolated from mice by using M-pluriBead Cell Separation kit (pluriSelect) following the procedure from the company and incubated with 10 μ g of mRNA-LNP in DC culture medium (RPMI 1640 supplemented with 20 ng/mL murine GM-CSF (R&D Systems), 10% FBS, 50 μ M 2-ME, 100

units/mL penicillin, and 100 µg/mL streptomycin) at 37°C for 48 h, then analyzed for MHC I and MHC II expression by flow cytometry.

[00160] Statistics and reproducibility. All data were presented as means ± standard error of the mean. The numbers of sample and replicates of experiments were shown as mentioned in the figure legends. Comparisons between groups were determined using Students t test. Differences were considered significant at *P < 0.001, **P < 0.05. All data were analyzed using GraphPad Prism 6 software.

EXAMPLE 1: Glycosylation of S protein affected antibody production

[00161] As a part of our efforts to identify possible conserved epitopes as targets for antibody development and next-generation vaccine design, and for the design of universal vaccine with broadly protective immune responses, we performed the S protein mutation analysis from the 218,516 available sequences of SARS-CoV-2 S protein. The S protein has 1,273 amino acids, and among the 218,516 sequences analyzed, there are 1,149 variable amino acid positions, including 613 in the S1 domain (672 amino acids), 524 in the S2 domain (588 amino acids), and 134 in the RBD (152 amino acid); however, the mutation rate is less than 0.1% at 1,076 sites while more than 0.1% at 73 sites. The conserved sequences can be found in the S1, S2 and the RBD regions, and the longest one is from R983-I1013 near the HR1 domain in the S2 region. All the 22 N-glycosylation sites are highly conserved among the SARS-CoV-2 variants. Further analysis of more sequences (about 6 million) show a similar distribution of conserved epitopes, 7 of which are in RBD and 5 in HR2 and 10 of the conserved epitopes are shielded by glycans (Fig. 1). We believe that removal of glycan shields on viral surface glycoproteins to expose more conserved epitopes is a very effective and general approach for vaccine design against SARS-CoV-2. Because the single GlcNAc residue linked to Asn is the minimum component of the N-glycan required for glycoprotein folding and stabilization, it is therefore postulated that trimming of N-glycans to leave a single GlcNAc on SARS-CoV-2 S protein will not affect its folding but will facilitate the maximum exposure of protein backbone to elicit robust and protein specific immune response while maintaining its structural integrity.

[00162] Based on our preliminary results, fully glycosylated (unmodified) and mono-GlcNAc decorated (on all N-glycosites) full length S, S1, S2, and RBD are used as immunogens for immunization studies. We produce S, S1, S2, and RBD subunits with retention of essential glycosites found in our preliminary study as described above and deletion of specific glycosites and their mono-GlcNAc decorated variants as immunogens for immunization. The antisera are tested for their interaction with representative S protein variants and neutralization activity against

pseudovirus-mediated infection, and those with broadly protective activities are further investigated including epitope mapping and adjuvant effect on CD4⁺ and CD8⁺ T-cell responses. The mono-GlcNAc decorated variants are made by removing the heterogeneous glycan layer on the *N*-glycosites of full-length S, S1, S2, and RBD that are produced using the more versatile and well demonstrated CHO, HEK293 or the Gnt1-deficient HEK293 cell line expression system. The glycans of the S protein expression in these cell lines can be trimmed with endoglycosidases to generate the desired protein with mono-GlcNAc at all the *N*-glycosylation sites and that from the latter (Gnt1-deficient HEK293) are high mannose types and can be digested using endoglycosidase H (Endo-H) to generate the desired protein with mono-GlcNAc at all the *N*-glycosylation sites. Since O-glycans are important for viral entry, no modification is carried out; but they can be trimmed with cocktails of exoglycosidases if necessary. This mono-GlcNAc decorated full length and truncated S proteins are studied regarding their structural integrity. Immunogens containing fully glycosylated and mono-GlcNAc proteins as well as the glycosite-engineered S protein (by replacing Asn with Gln as shown in the reverse genetics study) are used for mice immunization to identify antibodies that target the various domains on S protein with broad neutralization activity. The specificity of serum antibodies are checked by fully and mono-GlcNAc decorated as well as the glycosite-engineered S protein and its truncated forms. In addition, an array of synthetic peptides with or without mono-GlcNAc decorated or glycopeptides obtained from protease digestion of the mono-GlcNAc decorated S protein are used to study the binding specificity and the CD8⁺ T-cell response in transgenic mice with humanized ACE2 receptor. The immunized mice sera are further evaluated for the neutralization activity using pseudovirus neutralization assays developed in our lab. Fig. 2 shows the identification of N- and O-glycosites and mutations in variants. All 24 glycosites are highly conserved among 6 million S protein sequences.

[00163] S protein is frequently mutated and highly glycosylated with 22 N- and 2 O-glycosites (2 N- and 2 O-glycosites in RBD, and 6 N-glycosites in S2) to evade host immune response (Fig. 16). To study whether and how the mRNA vaccine of S protein with mutated glycosites affected S protein expression and immune response, we mutated multiple N-glycosites (from N to Q in the N-X-S/T sequon) and O-glycosites (from S/T to A) in the RBD or the S2 region of mRNA, respectively. After confirming the expression of the variant prefusion S protein in HEK293 cell line (Fig. 3), the mRNA that encoded S protein or S protein with mutation of glycosites, was encapsulated in LNP to form mRNA-LNP for immunization in mice. Sera from mice immunized by the mRNA with all S2 N-glycosites mutated (S-(deg-S2)) or with all S2 N-glycosites except N-1194 mutated (S-(S2-1194)), showed less IgG titer against fully glycosylated

WT S protein (Fig. 4A), S2 (Fig. 4B), RBD (Fig. 4C) or deglycosylated S protein (Fig. 4D), but had higher IgG titer against deglycosylated S2 antigen (Fig. 4E) in ELISA assay as compared to the unmodified mRNA. However, the mice immunized by the mRNA with all RBD glycosites deleted (S-(deg-RBD)) elicited a slightly less IgG titer against the fully glycosylated RBD, but higher IgG titer to recognize the deglycosylated RBD antigen (Fig. 4F), suggesting that glycosylation on S protein affected the production of antibody and its binding specificity. In addition, immunization with S-(deg-RBD), S-(deg-S2) or S-(S2-1194) mRNA elicited a higher IgG titer against the alpha (Fig. 5A), beta (Fig. 5B), gamma (Fig. 5C) delta, and omicron variants (Fig. 5D), suggesting that the glycosylation of S protein regulates the specificity of antibodies generated by the mRNA vaccine. To analyze the effect of glycosite mutation on the neutralization activity of antibodies generated from immunized mice, the pseudovirus neutralization assay was performed and the result showed that the mRNA vaccine with deletion of glycosites in RBD or S2 generated antibodies with reduced neutralization activity against WT pseudovirus (Fig. 5), but with better neutralization activity against the four variants of concern than WT (Fig. 6 and Table 1). To further understand this observation, sequence analysis revealed that mutation of the glycosites in RBD or S2 exposed more conserved epitopes to elicit immune responses (Fig. 6), and that the RBD and the S2 domains contained most of highly conserved sequences in the S protein (7 in RBD, 5 in HR2 of S2). These results suggested that mutation of certain glycosites in the mRNA of S protein vaccine will affect the production of antibodies and their specificities as well as the immune responses.

[00164] DNA or RNA sequence of WT S (Wuhan strain) (from 5'-end to 3'-end):

atgttcgtcttctggtcctgctgcctctggtctcctcacagtgcgtcaatctgacaactcggactcagctgccacctgcttatactaataagcttc
accagaggcgtgtactatcctgacaaggtgttagaagctccgtgctgcactctacacaggatctgtttctgccattcttagcaacgtgacct
ggtccacgccatccacgtgagcggcaccatggcacaagcggctcgacaatcccgtgctgccttttaacgatggcgtgtacttcgcctct
accgagaagagcaacatcatcagaggctggatcttggcaccacactggactccaagacacagtctctgctgatcgtgaacaatgccacc
aacgtggtcatcaaggtgtgcgagttccagtttgaatgatccctcctggcgctgtactatcacaagaacaataagagctggatggagtc
gagtttagagtgtattctagcgccaacaactgcacattgagtacgtgagccagccttctgatggacctggagggcaagcagggcaattt
caagaacctgagggagttcgtgtttaagaatatcgacggctacttcaaatctactctaagcacacccccatcaacctggtgcgcgacctgc
ctcagggcttcagcgccctggagccctggtggtatgcctatcgcatcaacatcacccggttcagacactgctggccctgcacagaag
ctacctgacacccggcgactcctctagcggatggaccgcccgcgtgccgctactatgtggctacctccagccccggaccttctgct
gaagtacaacgagaatggcaccatcacagacgcagtggtgattgcgcctggacccctgagcgagacaaagtgtactgaagtcctttac
cgtggagaagggcatctatcagacatccaattcaggggtgcagccaacaggatctatcgtgcgcttcttaatatcacaacctgtgccatt
tggcgagggtgttaacgcaacccgcttcgccagcgtgtacgcctggaataggaagcggatcagcaactgcgtggccgactatagcgtgct
gtacaactccgcctctttagcacctttaagtgtatggcgtgtccccacaaagctgaatgacctgtgctttaccaacgtctacgccgattctt

tcgtgatcagggcgacgaggtgcgccagatcggccggcagacaggcaagatcgagactacaattataagctgccagacgatttc
accggctgctgatcgctggaacagcaacaatctggattcaaagtggcgggcaactacaattatctgtaccggctgttagaaagagca
atctgaagccctcgagaggacatctctacagaaatctaccaggccggcagcacccttgaatggcgtggagggtttaactgttatttc
ccactccagtcctacggctccagcccacaaacggcgtgggctatcagccttaccgctgggtggtgctgagctttgagctgctgcacggcc
cagcaacagtgtgcggcccaagaagtccaccaatctggtgaagaacaagtgcgtgaactcaactcaacggcctgaccggcacaggc
gtgctgaccgagtcacaagaagtctgcatctcagcagttcggcaggacatcgagataccacagacggcgtgcgcgaccaca
gaccttgagatcctggacatcacacctgctcttccggcggtgagcgtgatcacccggcaccaatacaagcaaccagggtggccgt
gctgtatcaggacgtgaattgtaccgaggtgccgtggctatccacgccgatcagctgacccaacatggcgggtgtacagcaccggctc
caacgtctccagacaagagccggatgctgatcgagcagagcacgtgaacaattcctatagtgacatcccaatggcgccggcat
ctgtgctcttaccagaccagacaaactctccagaagagccggagcgtggcctccagtctatcatcgctataccatgtcctgggc
gcccgaacagcgtggcctacttaacaatgacatcgccatcccaactcaaatctctgtgaccacagagatcctgcccgtgtccat
gaccaagacatctgtgactgcacaatgtatatctgtggcgattctaccgagtcagcaacctgctgctccagtacggcagctttgtacca
gctgaatagaccctgacaggcatcgccgtggagcaggataagaacacacaggaggtgttcgccagggtgaagcaaatctacaagacc
ccccctatcaaggactttggcggttcaatcttccagatcctgctgatccatccaagccttctaagcggagctttatcaggacgtgctgtt
caacaagggtgacctggccgatcgccgttcatcaagcagtatggcgattgctggcgacatcgagccaggacgtgatctgcgcc
agaagttaatggcctgaccgtgctgccaccctgctgacagatgagatgatcgacagtacacaagcgccctgctggccggcaccatca
catccgatggacctcgccgacaggagccgcccctccagatcccccttggcatgacagtgccctatagggttaacggcatcgccgtgaccc
agaatgtgctgtacgagaaccagaagctgatcgccaatcagtttaactccgccatcggaagatccaggacagcctgtcctctacagccag
cgccctgggcaagctccaggatgtggtgaatcagaacgcccaggccctgaataccctggtgaagcagctgagcagcaacttcggcgcc
atctctagcgtgctgaatgacatcctgagccggctggacaagggtggaggcagaggtgcagatcgaccggctgatcaccggccggctcca
gagcctccagacctatgtgacacagcagctgatcaggccgcccagatcaggccagcgccaatctggcagcaaccaagatgtccgag
tgctgtgctggccagctaaagaggtgacttttgggcaagggtatcacctgatgtcctccctcagctgtccccacacggcgtggtgttt
ctgcacgtgacctacgtgcccggcagagaagaacttcaccacagcccctgccatctgccacgatggcaaggcccactttccaaggga
ggcgctgttcgtgtcaacggcacccactggtttgtgacacagcgcaatttctacgagcccagatcatcaccacagacaacaccttcgtg
agcggcaactgtgacgtggtcatcgccatcgtgaacaatacctgtgatccactccagcccagctggacagctttaaggaggagctg
gataagtatttcaagaatcacacctcccctgacgtggatctggcgacatcagcggcatcaatgcctccgtggtgaacatccagaaggaga
tcgaccgctgaacgaggtggctaagaatctgaacgagagcctgatcgacctccaggagctgggcaagtatgagcagtagcatcaagtgg
ccctggtacatctggctgggttcatcgccggcctgatcgccatcgtgatggtgacctcatgctgtgctgtatgacatcctgctgttcttgcct
gaagggtgctgtagctgtggctcctgctgtaagtttgacgaggatgactctgaacctgtgctgaaggcgctgaagctgcattacacctaa
(SEQ ID NO: 1)

[00165] Protein sequence of WT S (Wuhan strain) (from N-terminus to C-terminus):

MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFS
NVTWFHAIHVS GTNGTKRFDNPVLPFNDGVYFASTKSNIIRGWIFGTTLD SKTQSL LIV
NNATNVVIK VCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMD

LEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQ
 TLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLS
 ETKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNKRRI
 SNCVADYSVLNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVQRQIAPGQTGK
 IADYNYKLPPDDFTGCVIAWNSNNLDSKVGGNYNLYRLFRKSNLKPFERDISTEIQAG
 STPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVLSFELLHAPATVCGPKKSTNLV
 KNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGG
 VSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGA
 EHVNNSEYCDIPIGAGICASYQTQTNSPRRARSVASQSIIAYTMSLGAENSVAYSNNISIAI
 PTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLQYGSFCTQLNRALTGIAVEQD
 KNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSPKPSKRSFIEDLLFNKVTLADAGFIKQ
 YGDCLGDIAARDLCAQKFNGLTVLPPLLDEMIAQYTSALLAGTITSGWTFGAGAALQ
 IPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNESAIGKIQDSLSTASALGKLQDVVNQ
 NAQALNTLVKQLSSNFGAISSVLNDILSRLDPPAEVQIDRLITGRLQSLQTYVTQQLIRA
 AEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSPQSAPHGVVFLHVTYVPAQEK
 NFTTAPAICHGDKAHFPREGVFVSNGTHWFVTQRNFYEPQIITDNTFVSGNCDVVIGIV
 NNTVYDPLQPELDSFKEELDKYFKNHTSPDVLGDISGINASVVNIQKEIDRLNEVAKNL
 NESLIDLQELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCCLKGCCSCGSC
 CKFDEDDSEPVLKGVKLHYT (SEQ ID NO: 2)

[00166] DNA or RNA sequence of WT S (Delta strain) (from 5'-end to 3'-end):

atgttcgtcttctggtcctctgctgcctctggtctcctcacagtgcgtcaatctgcgaactcggactcagctgccacctgcttatactaagcttc
 accagaggcgtgtactatcctgacaaggtgttagaagctccgtgctgcactctacacaggatctgttttccattcttagcaacgtgacct
 ggttcacgccatccacgtgagcggcaccaatggcacaagcgggtcgacaatcccgtgctgcctttaacgatggcgtgtacttcgctct
 atcgagaagagcaacatcatcagaggctggatcttggcaccacactggactccaagacacagtctctgctgatcgtgaacaatgccacca
 acgtggtcatcaaggtgtgcgagttccagtttgaatgatccctcctggacgtgtactatcacaagaacaataagagctggatggagtcgg
 gagtgtattctagcgccaacaactgcacattgagtagctgagccagccttctctgatggacctggagggcaagcagggcaattcaagaa
 cctgaggaggtcgtgttaagaatatcgacggctacttcaaatctactctaagcacacccccatcaacctggtgcgcgacctgcctcagg
 gcttcagcgccctggagcccctggtggatctgcctatcggcataacatcacccggttcagacactgctggccctgcacagaagctacct
 gacacccggcgactcctctagcggatggaccgcccgcgtgccgctactatgtgggctacctccagccccggaccttctgctgaagta
 caacgagaatggcaccatcacagacgcagtgattgcgcctggacccctgagcagagacaaagtgtactgaagtcctttaccgtgga
 gaagggcatctatcagacatccaatttcagggtgcagccaaccgagtcctatcgtgcgctttcctaataatcacaaacctgtcccatttggcga
 ggtgttaacgcaacccgcttcgccagcgtgtacgcctggaataggaagcggatcagcaactgcgtggccgactatagcgtgctgtacaa
 ctccgcctcttccagacatttaagtgtatggcgtgtccccacaaagctgaatgacctgtgctttaccaacctctacgccgattcttctgtgat
 caggggacgaggtgcgcccagatgccccggccagacaggcaagatcgagactacaattataagctgcccagacgatttcaccggc

tgcgtgatcgctggaacagcaacaatctggattccaaagtggcggaactacaattatcggtaccggctgttagaaagagcaatctga
 agcccttcgagaggacatctctacagaaatctaccaggccggcagcaagccttgcaatggcgtggagggttaactgttattccactc
 cagtctacggcttcagcccacaaacggcgtgggctatcagccttaccgctgggtgctgagctttgagctgctgcacgcccagca
 acagtgtgcggcccaagaagtccaccaatctggtgaagaacaagtgcgtgaactcaactcaacggcctgaccggcacaggcgtgct
 gaccgagccaacaagaagtctcgtccatttcagcagttcggcaggacatcgagataccacagacggcgtgcgcgaccacagacc
 tggagatcctggacatcacacctgctcttcggcggtgagcgtgatcacccggcaccaatacaagcaaccagggtggcgtgctgt
 atcaggggcgtgaattgtaccagggtccccgtggctatccacggcatcagctgaccccaacatggcgggtgtacagcaccggctccaa
 gtcttcagacaagagccggatgcctgatcggagcagagcacgtgaacaattcctatgagtgcgacatcccaatcgcgccggcatctgt
 gcctctaccagaccagacaaactctcgcagaagagccggagcgtggcctcccagctatcatcgctatacatgtccctggcgcc
 gagaacagcgtggcctacttaacaatagcatcgccatcccaaccaacttcacaatctctgtgaccacagagatcctgcccgtgtccatgac
 caagacatctgtggatgcacaatgtatatctgtggcgttctaccgagtgacgaacctgctgctccagtagcgcagctttgtaccagct
 gaatagagccctgacaggcatcgccgtggagcaggataagaacacacaggaggtgttcgccagggtgaagcaaatctacaagaccccc
 cctatcaaggactttggcggtcaattttccagatcctgctgatccatcaagccttcaagcggagctttatcgaggacctgtgttaa
 caaggtgacctggccgatcgccggttcataagcagtatggcgattgcctggcgacatcgagccaggacctgatctgcgccaga
 agtttaaggcctgacctgctgccacctgctgacagatgagatgatcgacagtacacaagcgccctgctggccggcaccatcacatc
 cggatggaccttcggcgaggagccgcccctccagatcccccttgccatgcagatggcctataggttaacggcatcgccgtgaccagaa
 tgtgtgtacgagaaccagaagctgatcgccaatcagtttaactccgcatcggaagatccaggacagcctgtcctctacagccagcgcc
 ctgggcaagctccagaatgtgtgaatcagaacgcccaggccctgaataacctggtgaagcagctgagcagcaacttcggcgccatctct
 agcgtgctgaatgacatcctgagccggctggacaaggtggaggcagaggtgcagatcgaccggctgatcaccggccggctccagagc
 ctccagacctatgtgacacagcagctgatcaggccgcccagatcaggccagcgccaatctggcagcaaccaagatgtccgagtgct
 gctggccagcttaagagagtggtgactttgtggcaagggtatcacctgatgtccttcctcagctctgccccacagcggtggtgtttctgca
 cgtgacctacgtgcccggcaggagaagaacttcaccacagcccctgccatctgccacgatggcaaggcccactttccaaggaggggc
 gtgttcgtgtccaacggcaccactggtttgtgacacagcgcaatttctacagccccagatcatcaccagacaacacctctgtgagcg
 gcaactgtgagctggtcatcgcatcgtgaacaataccgtgtatgatccactccagcccagctggacagctttaaggaggagctggataa
 gtatttcaagaatcacacctccccctgacgtggatctggcgacatcagcgcatcaatgcctccgtggtgaacatccagaaggagatcgac
 cgctgaacgaggtggctaagaatctgaacgagagcctgatcgacctccaggagctgggcaagtatgagcagtagcatcaagtggccctg
 gtacatctggctgggcttcacgcccgtgatcgccatcgtgatggtgacctcatgctgtgctgtatgacatcctgctgttcttgctgaag
 ggctgctgtagctgtggctcctgctgtaagttgacgaggatgactctgaacctgtgctgaaggcgctgaagctgcattacacctaa
 (SEQ ID NO: 15)

[00167] Protein sequence of WT S (Delta strain) (from N-terminus to C-terminus):

MFVFLVLLPLVSSQCVNLRTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFS
 NVTWFHAIHVSNGTKRFDNPVLPFNDGVYFASIEKSNIIRGWIFGTTLDSTQSLIV
 NNATNVVVKVCEVFQCNDFLDVYYHKNNKSWMESGVYSSANNCTFEYVSQPFLMDL
 EGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQT

LLALHRSYLTPGDSSSGWTAGAAAYVGYLQPRTFLLKYNENGTITDAVDCALDPLSE
 TKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRRKRIS
 NCVADYSVLNYSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVQRQIAPGQTGKI
 ADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYRYRLFRKSNLKPFERDISTEIQAG
 SKPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVLSFELLHAPATVCGPKKSTNLV
 KNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGG
 VSVITPGTNTSNQVAVLYQGVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGA
 EHVNNSYECDIPIGAGICASYQTQTSNRRRARSVASQSIIAYTMSLGAENSVAYSNNNSIAI
 PTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLQYGSFCTQLNRALTGIAVEQD
 KNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQ
 YGDCLGDIAARDLCAQKFNGLTVLPPLLDEMIAQYTSALLAGTITSGWTFGAGAALQ
 IPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSTASALGKLQNVVNQ
 NAQALNTLVKQLSSNFGAIVSVLNDILSRDLKVEAEVQIDRLITGRLQSLQTYVTQQLIR
 AAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLSFPQSAPHGVVFLHVTYVPAQE
 KNFTTAPAICHGDKAHFPREGVFSNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGI
 VNNTVYDPLQPELDSFKEELDKYFKNHTSPDVLDGDISGINASVVNIQKEIDRLNEVAK
 NLNESLIDLQELGKYEYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCCLKGCCSC
 GSCKFDEDDSEPVLKGVKLHYT (SEQ ID NO: 16)

[00168] DNA or RNA sequence of WT S (Wuhan strain S-2P strain) (from 5'-end to 3'-end):

atgttcgtcttctggtcctgctgcctctggtctcctcacagtgcgtcaatctgacaactcggactcagctgccacctgcttataactaatagcttc
 accagaggcgtgtactatcctgacaaggtgttagaagctccgtgctgcactctacacaggatctgtttctgccattcttagcaacgtgacct
 ggttcacgccatccacgtgagcggcaccatggcacaagcggctcgacaatcccgtgctgcctttaacgatggcgtgtacttcgctct
 accgagaagagcaacatcatcagaggctggatcttggcaccacactggactccaagacacagtctctgctgatcgtgaacaatgccacc
 aacgtggtcatcaaggtgtgcgagttccagttttgtaatatgaccttctggcgctgtactatcacaagaacaataagactggatggagtc
 gaggtttagagtgtattctagcgccaacaactgcacatttgagtacgtgagccagccttctgatggacctggagggcaagcagggcaattt
 caagaacctgagggagttcgtgttaagaatatcgacggctacttcaaatctacttaagcacacccccatcaacctggtgcgcgacctgc
 ctgagggcttcagcgccctggagcccctggtgatctgcctatcgcatcaacatcacccggttcagacactgctggccctgcacagaag
 ctacctgacacccggcgactcctctagcggatggaccgcccgcgtgccgctactatgtggctacctccagccccggaccttctgct
 gaagtacaacgagaatggcaccatcacagacgcagtggttgcgcctggacccctgagcgagacaaaagtgtactgaagtcccttac
 cgtggagaagggcatctatcagacatccaatttcagggtgcagccaaccgagtgctatcgctgcgtttctaatatcacaacctgtgccatt
 tggcgagggtgtcaacgcaaccgcttcgagcgtgtacgcctggaataggaagcggatcagcaactgcgtggccgactatagcgtgct
 gtacaactccgcctcttctcagcacctttaaagtgtatggcgtgtccccacaaagctgaatgacctgtgctttaccaacgtctacgcgattctt
 tcgtgatcagggcgacgaggtgcgccagatcgccccggcagacaggaagatcgagactacaattataagctgccagacgatttc

accggctgctgatcgcttggaacagcaacaatctggattcacaagtggcgcgcaactacaattatctgtaccggctgttagaaagagca
 atctgaagcccttcgagaggacatctctacagaaatctaccaggccggcagcacccttgcaatggcgtggaggctttaactgttatttc
 ccactccagctctacggcttcagcccacaaacggcgtgggctatcagccttacggcgtgggtgctgagctttgagctgctgcacggcc
 cagcaacagtgtgcgcccccaagaagtccaccaatctggtgaagaacaagtgcgtgaactcaactcaacggcctgaccggcacaggc
 gtgctgaccgagtgcaacaagaattcctgccatttcagcagttcggcaggacatcgagataccacagacggcgtgcgcgaccaca
 gaccctggagatcctggacatcacacctgctcttcggcgcggtgagcgtgatcacccggcaccaatacaagcaaccagggtggcgt
 gctgtatcaggacgtgaattgtaccgaggtgcccgtggtatccacgcccgtacgtgaccccaacatggcggtgtacagaccggctc
 caacgtcttcagacaagagccggatgcctgatcggagcagagcacgtgaacaattcctatgagtgcgacatcccaatcgcgccggcat
 ctgtgctcttaccagaccagacaaactctccagaagagccggagcgtggcctccagtctatcatcgctatacatgtcctgggc
 gccgagaacagcgtggcctactctaacaatgacatcgccatcccaaccaactcacaatctctgtgaccacagagatcctgcccgtgccat
 gaccaagacatctgtggactgcacaatgtatatctgtggcgattctaccgagtgagcaacctgctgctccagtagcgagctttgtacca
 gctgaatagagccctgacaggcatcgccgtggagcaggataagaacacacaggaggtgtcggcagggtgaagcaaattacaagacc
 cccctatcaaggactttggcggttcaattttccagatcctgctgatccatccaagccttcaagcggagctttatcaggacgtgctgtt
 caacaaggtagacctggccgatgccggctcatcaagcagtagtgcgattgcctggcgacatcgagccaggacgtgatctgcggcc
 agaagtttaattggcctgaccgtgctgccaccctgctgacagatgagatgatcgacagtagacaaagcgccctgctggcgccaccatca
 catccggatggaccttcggcgaggagccgcccctccagatcccccttggcatgagatggcctatagggttaacggcatcgcggtgacc
 agaatgtgctgtacgagaaccagaagctgatcgccaatcagtttaactccgcatcggaagatccaggacagcctgtcctctacagccag
 cgccctgggcaagctccaggatgtggtgaatcagaacgcccaggccctgaataccctggtgaagcagctgagcagcaacttcggcgcc
 atctctagcgtgctgaatgacatcctgagccggctggacccgagcagaggtgcagatcgaccggctgatcacggcgccgctcc
 agagcctccagacctatgtgacacagcagctgatcagggccgagatcagggccagcgccaatctggcagcaaccaagatgtccga
 gtgctgctgggccagtctaagagagtggaactttgtggcaagggtatcacctgatgtccttcctcagctgtccccacacggcgtggtgtt
 tctgcacgtgacctacgtgcccggcaggagaagaattccaccagcccctgccatctgccagatggcaaggcccactttcaaggga
 gggcgtgttcgtgtcaacggcaccactgtgttgacacagcgcaatttctacgagcccagatcatcaccacagacaacaccttcgtg
 agcggcaactgtgacgtggtcatcgccatcgtgaacaataccgtgtatgatccactccagcccagctggacagctttaaggaggagctg
 gataagtattcaagaatcacacctcccctgacgtggatctggcgacatcagcggcatcaatgcctccgtggtgaacatccagaaggaga
 tcgaccgctgaacgaggtggctaagaatctgaacgagagcctgatcgacctcaggagctgggcaagtatgagcagtagcatcaagtgg
 ccctggtacatctggctgggcttcacgcccgtgatcgccatcgtgatggtgacctcatgctgtgctgtatgacatcctgctgttcttgcct
 gaagggtgctgtagctgtggctcctgctgtaagtttgacgaggatgactctgaacctgtgctgaaggcgctgaagctgcattacacctaa
 (SEQ ID NO: 17)

[00169] **Protein sequence of WT S (Wuhan strain S-2P strain) (from N-terminus to C-terminus):**

MFVFLVLLPLVSSQCVNL TTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFS
 NVTWFHAIHVS GTNGTKRFDNPVLPFNDGVYFASTKSNIRGWIFGTTLD SKTQSLIV
 NNATNVVIK VCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMD

LEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQ
 TLLALHRSYLTTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLS
 ETKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNKRRI
 SNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVQRQIAPGQTGK
 IADYNYKLPPDDFTGCVIAWNSNNLDSKVGGNYNLYRLFRKSNLKPFERDISTEIQAG
 STPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVLSFELLHAPATVCGPKKSTNLV
 KNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGG
 VSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGA
 EHVNNSEYCDIPIGAGICASYQTQTNSPRRARSVASQSIIAYTMSLGAENSVAYSNNISIAI
 PTNFTISVTTEILPVSMKTSTVDCTMYICGDSTECSNLLQYGSFCTQLNRALTGIAVEQD
 KNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSPKPSKRSFIEDLLFNKVTLADAGFIKQ
 YGDCLGDIAARDLCAQKFNGLTVLPLLDEMIAQYTSALLAGTITSGWTFGAGAALQ
 IPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNESAIGKIQDSLSTASALGKLQDVVNQ
 NAQALNTLVKQLSSNFGAISSVLNDILSRLD[PP]EAEVQIDRLITGRLQSLQTYVTQQLIRA
 AEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSEFPQSAPHGVVFLHVTYVPAQEK
 NFTTAPAICHGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITDNTFVSGNCDVVIGIV
 NNTVYDPLQPELDSFKEELDKYFKNHTSPDVLGDISGINASVVNIQKEIDRLNEVAKNL
 NESLIDLQELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCCLKGCCSCGSC
 CKFDEDDSEPVLKGVKLHYT (SEQ ID NO: 18)

[00170] DNA or RNA sequence of WT S (Delta S-2P strain) (from 5'-end to 3'-end):

atgttcgtcttctggtcctgctgcctctggtctcctcacagtgcgtcaatctgcgaactcggactcagctgccacctgcttatactaataagcttc
 accagaggcgtgtactatcctgacaaggtgttagaagctcctgctgcactctacacaggatctgtttctgccattctttagcaacgtgacct
 ggttcacgccatccacgtgagcggcaccatggcacaagcgggtcgacaatcccgctgctgccttttaacgatggcgtgtacttcgctct
 atcgagaagagcaacatcatcagaggctggtatcttggcaccacactggactccaagacacagtctctgctgatcgtgaacaatgccacca
 acgtggtcatcaagggtgctgagttccagttttgtaatgatccctcctggacgtgtactatcacaagaacaataagagctggatggagtccg
 gagtgtattctagcgccaacaactgcacatttgagtacgtgagccagcctttcctgatggacctggagggaagcagggaattcaagaa
 cctgaggggagttcgtgtttaagaataatcgacggctacttcaaaatctactctaagcacacccccatcaacctggcgcgacctgcctcagg
 gcttcagcgccctggagcccctgggtgatctgcctatcggcataacatcacccggttcagacactgctggccctgcacagaagctacct
 gacacccggcgactcctctagcggatggaccgcccggcgtgcccctactatgtgggctacctccagccccggacctcctgctgaagta
 caacgagaatggcaccatcacagacgcagtggtgattgcgccctggacccctgagcgagacaaagtgtactgaagtcctttaccgtgga
 gaagggcatctatcagacatccaatttcagggtgcagccaaccgagtcctatcgtgcgctttcctaataatcacaacctgtgccatttggcga
 ggtgttaacgcaacccgcttcgccagcgtgtacgcctggaataggaagcggatcagcaactgcgtggccgactatagcgtgctgtacaa
 ctccgcctcttcagcacctttaagtgtatggcgtgtccccacaaagctgaatgacctgtgctttaccaacgtctacgccgattcttcgtgat
 cagggggcgacgaggtgctgagatcgccccggccagacagggaagatcgagactacaattataagctgccagacgatttcaccggc

tgcgtgatcgctggaacagcaacaatctggattccaaagtggcggaactacaattatcggtaccggctgttagaaagagcaatctga
 agcccttcgagaggacatctctacagaaatctaccaggccggcagcaagccttgcaatggcgtggagggttaactgttattccactc
 cagtctacggcttcagcccacaaacggcgtgggctatcagccttaccgctgggtgctgagctttgagctgctgcacgcccagca
 acagtgtgcggcccaagaagtccaccaatctggtgaagaacaagtgcgtgaactcaactcaacggcctgaccggcacaggcgtgct
 gaccgagtccaacaagaagtctcgtccatttcagcagttcggcaggacatcgagataccacagacggcgtgcgcgaccacagacc
 tggagatcctggacatcacacctgctcttcggcggtgagcgtgatcacccggcaccaatacaagcaaccagggtggcgtgctgt
 atcaggggcgtgaattgtaccagggtccccgtggctatccacggcatcagctgaccccaacatggcgggtgtacagcaccggctccaa
 gtcttcagacaagagccggatgcctgatcggagcagagcacgtgaacaattcctatgagtgcgacatcccaatcgcgccggcgtctgt
 gcctcttaccagaccagacaaactctcgcagaagagccggagcgtggcctcccagctatcatcgctatacatgtccctggcgcc
 gagaacagcgtggcctacttaacaatagcatcgccatcccaaccaacttcacaatctctgtgaccacagagatcctgcccgtgtccatgac
 caagacatctgtggatgcacaatgtatatctgtggcgattctaccgagtgcagcaacctgctgctccagtagggcagctttgtaccagct
 gaatagagccctgacaggcatcgccgtggagcaggataagaacacacaggaggtgttcgccagggtgaagcaaatctacaagaccccc
 cctatcaaggactttggcggtcctaattttccagatcctgctgatccatccaagccttcaagcggagctttatcgaggacctgtgttaa
 caaggtgacctggccgatcgccggttcataagcagtatggcgattgcctggcgacatcgagccaggacctgatctgcgccaga
 agtttaaggcctgacctgctgccacctgctgacagatgagatgatcgacagtacacaagcgcctgctggccggcaccatcacatc
 cggatggaccttcggcgaggagccgcccctccagatcccccttgccatgcagatggcctatagggttaacggcatcgccgtgaccagaa
 tgtgtgtacgagaaccagaagctgatcgccaatcagtttaactccgcatcggaagatccaggacagcctgtccttacagccagcgcc
 ctgggcaagctccagaatgtgtgaatcagaacgcccaggccctgaataacctgtgaagcagctgagcagcaacttcggcgccatctct
 agcgtgctgaatgacatcctgagccggctggacccgagcagaggtgcagatcgaccggctgatcaccggccggctccagagc
 ctccagacctatgtgacacagcagctgatcaggccggcggagatcaggccagcgccaatctggcagcaaccaagatgtccgagtgcgt
 gctgggcccagtctaagagagtggaactttgtggcaagggtatcacctgatgtccttccctcagctctgccccacacggcgtggtgttctgca
 cgtgacctacgtgcccggccaggagaagaacttcaccacagcccctgcccctgacagtggaaggcccactttcaaggaggggc
 gtgtctgtgtccaacggcaccactgtgtgtgacacagcgcaatttctacagccccagatcatcaccacagacaacaccttctgtgagcg
 gcaactgtgacgtggtcatcgccatcgtgaacaataccgtgtatgatccactccagcccagctggacagctttaaggaggagctggataa
 gtatttcaagaatcacacctccccctgacgtggatctggcgacatcagcggcatcaatgcctccgtggtgaacatccagaaggagatcgac
 cgccgaacgaggtggctaagaatctgaacgagagcctgatcgacctccaggagctgggcaagtatgagcagtagcatcaagtggccctg
 gtacatctggctgggcttcatcgccggcctgatcgccatcgtgatggtgacctcatgctgtgctgtatgacatcctgctgttcttgctgaag
 ggctgctgtagctgtggctcctgctgtaagtttgacgaggatgactctgaacctgtgctgaaggcgctgaagctgcattacacctaa
 (SEQ ID NO: 19)

[00171] Protein sequence of WT S (Delta strain S-2P strain) (from N-terminus to C-terminus):

MFVFLVLLPLVSSQCVNLRTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFS
 NVTWFHAIHVS GTNGTKRFDNPVLPFNDGVYFASIEKSNIIRGWIFGTTLDSKTQSLIV
 NNATNVVIK VCEFQFCNDPFLDVYYHKNNKSWMESGVYSSANNCTFEYVSQPFLMDL

EGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQT
 LLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSE
 TKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRRKRIS
 NCVADYSVLVNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVQRQIAPGQTGKI
 ADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNRYRLFRKSNLKPFERDISTEIQAG
 SKPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVLSFELLHAPATVCGPKKSTNLV
 KNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGG
 VSVITPGTNTSNQVAVLYQGVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGA
 EHVNNSEYCDIPIGAGICASYQTQTSNRRRRARSVASQSIIAYTMSLGAENSVAYSNNNSIAI
 PTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLQYGSFCTQLNRALTGIAVEQD
 KNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQ
 YGDCLGDIAARDLCAQKFNGLTVLPPLLDEMIAQYTSALLAGTITSGWTFGAGAALQ
 IPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSTASALGKLQNVVNQ
 NAQALNTLVKQLSSNFGAISSVLNDILSRLD[PP]EAEVQIDRLITGRLQSLQTYVTQQLIRA
 AEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEK
 NFTTAPAICHGKAHFPREGVSVSNGTHWFTQRNFYEPQIITDNTFVSGNCDVVIGIV
 NNTVYDPLQPELDSFKEELDKYFKNHTSPDVLGDISGINASVVNIQKEIDRLNEVAKNL
 NESLIDLQELGKYEQYIKWPWYIWLGFIAGLIAIVMTIMLCCMTSCCCLKGCCSCGSC
 CKFDEDDSEPVLKGVKLHYT (SEQ ID NO: 20)

[00172] DNA or RNA sequence of S-(deg-RBD) for Wuhan strain (from 5'-end to 3'-end):

atgttcgtcttctggtcctgctgcctctggtctcctcacagtgcgtcaatcgacaactcggactcagctgccacctgcttataactaatagcttc
 accagaggcgtgtactatcctgacaaggtgttagaagctcgtgctgcactctacacaggatctgtttctgccattctttagcaacgtgacct
 ggttcacgccatccacgtgagcggcacaatggcacaagcggctcgacaatcccgtgctgccttttaacgatggcgtgtacttcgctct
 accgagaagagcaacatcatcagaggctggatcttggcaccacactggactccaagacacagtctctgctgatcgtaacaatgccacc
 aacgtggatcatcaaggtgtgcgagttccagttttgtaatgatccctcctggcgctgtactatcacaagaacaataagagctggatggagtc
 gagtttagagtgtattctagcgccaacaactgcacattgagtagtgagccagccttctgatggacctggagggaagcagggaattt
 caagaacctgagggagttcgtgtttaagaatatcgacggctacttcaaatctactctaagcacacccccatcaacctggtgcgcgacctgc
 ctgagggcttcagcgccctggagcccctggtggatctgcctatcgcatcaacatcacccgggttcagacactgctggccctgcacagaag
 ctacctgacacccggcgactcctctagcggatggaccgcccgcgtgccgcctactatgtgggctacctccagccccggaccttctgct
 gaagtacaacgagaatggcaccatcacagacgcagtggtgctgcctggacccctgagcgagacaaagtgtacactgaagtcctttac
 cgtggagaagggcatctatcagacatccaatttcagggtgcagcca[ccgag]gcatctgctgcttct[gaat]cacaacctgtgcc
 atttggcgaggtgtc[ga]gcaacccgcttgcgcagcgtgtacgcctggaataggaagcggatcagcaactgcgtggccgactatagcgt
 gctgtacaactccgccttcttcagcaccttaagtctatggcgtgtccccacaaagctgaatgacctgtgctttaccaacgtctacgccgat

tctttcgtgatcaggggacgaggtgcccagatcgccccggccagacaggcaagatcgagactacaattataagctgccagacga
 tttcaccggctgctgatcgctggaacagcaacaatctggattccaaagtggcggaactacaattatctgtaccggctgttagaaga
 gcaatctgaagcccttcgagaggacatctctacagaatctaccaggccggcagcacccttgcaatggcgtggagggtttaactgtta
 ttccactccagtcctacggcttcagcccacaaacggcgtgggctatcagcctaccgctgggtgctgagctttgagctgtgcacg
 cccagcaacagtgtgcccccaagaagtcaccaatctggtagaacaagtgcgtgaactcaactcaacggcctgaccggcaca
 ggctgtgacaggatccaacaagaagttcctgccatttcagcagttcggcaggacatcgagataccacagacggcgtgcgcgaccc
 acagaccctggagatcctggacatcacacctgctcttcggcggtgagcgtgatcacaccggcaccataaagcaaccaggtgg
 ccgtgctgtatcaggacgtgaattgtaccgaggtgcccgtggctatccacggcagcgtgaccccaacatggcggtgtacagaccg
 gctcaacgtcttcagacaagagccggatgctgatcgagcagagcagcgtgaacaattctatgagtcgacatcccaatcgcgccg
 gcatctgtcctctaccagaccagacaaactctccagaagagccggagcgtggcctccagctctatcatgcctatacatgtcctg
 ggcgcgagaacagcgtggcctacttaacaatagcatcgccatcccaacttcacaatctctgtgaccacagagatcctgcccgtg
 ccatgaccaagacatctgtgactgcacaatgtatatctgtggcgattctaccgagtgagcaacctgctgctccagtacggcagctttgta
 cccagctgaatagagccctgacaggcatcgccgtggagcaggataagaacacagaggtgttcgccaggtgaagcaaatctaca
 gacccccctatcaaggactttggcggttcaattttccagatcctgctgatccatcaagccttctaagcggagctttatcaggacctg
 ctgttcaacaagtgacccctggcggttcataagcagtatggcgattgctggcgacatcgagccaggacctgatctgc
 gccagaagttaatggcctgaccgtgctgccaccctgctgacagatgagatgatcgacagtacacaagcgccctgctggccggcacc
 atcacatccggatggaccttcggcgaggagccgctccagatccctttgccatgcagatggcctataggttcaacggcatcgcggtga
 cccagaatgtgctgtacgagaaccagaagctgatcgcaatcagtttaactccgcatcggaagatccaggacagcctgtcctctacagc
 cagcgccctgggcaagctccagatgtggtgaatcagaacgccagccctgaatacctggtgaagcagctgagcagcaacttcggc
 gccatctctagcgtgctgaatgacatcctgagccggctggacaaggtggaggcagaggtgcagatcgaccggctgatcaccggccggct
 ccagagcctccagacatgtgacacagcagctgatcaggccggcagatcaggccagcgcaatctggcagcaaccaagatgtcc
 gagtgctgctgggcccagctaaagagagtggaactttgtggcaagggtatcacctgatgtccttccctcagctgtccccacacggcgtggt
 gtttctgcacgtgacctacgtgcccggcaggagaagaacttcaccacagcccctgcatctgccagatggcaaggcccactttcaag
 ggaggcgctgttctgttccaacggcaccactggtttgtgacacagcgcaatttctacgagccccagatcatcaccagacacaacacctc
 gtgagcggcaactgtgacgtggtcatcggcacgtgaacaatacctgtatgatccactccagcccagctggacagctttaaggaggag
 ctggataagtatttcaagaatcacacctcccctgacgtggatctggcgacatcagcgcatcaatgcctccgtggtgaacatccagaagg
 agatcgaccgcctgaacgaggtggctaagaatctgaacgagagcctgatcgacctccaggagctgggcaagtatgagcagctacatcaag
 tggccctggtacatctggctgggcttcacgcccgtgatcgccatcgatggtgacctcatgctgtgctgtatgacatcctgctgttctt
 gcctgaagggtgctgtagctgtggctcctgctgtaagtttgacgaggatgactctgaacctgtgctgaaggcggtgaagctgcattacac
 taa (SEQ ID NO: 3)

[00173] Protein sequence of S-(deg-RBD) for Wuhan strain (from N-terminus to C-terminus):

MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFS
 NVTWFHAIHVS GTNGTKRFDNPVLPFNDGVYFASTKSNIIRGWIFGTTLD SKTQSLLIIV

NNATNVVIKVFCEFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMD
 LEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQ
 TLLALHRSYLTTPGDSSSGWTAGAAAYVGYLQPRTFLLKYNENGTITDAVDCALDPLS
 ETKCTLKSFTVEKGIYQTSNFRVQP[A][E]IVRFP[Q]ITNLCPFGEVF[Q]ATRFASVYAWNRRK
 RISNCVADYSVLVNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVQRQIAPGQTG
 KIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNLYRLFRKSNLKPFERDISTEYQA
 GSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVLSFELLHAPATVCGPKKSTNL
 VKNKCVMNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFG
 GVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIG
 AEHVNNSYECDIPIGAGICASYQTQTNSPRRARSVASQSIIAYTMSLGAENSVAYSNNSI
 AIPTNFTISVTTEILPVSMTKTSVDCCTMYICGDSTECNLLLQYGSFCTQLNRALTGIAVE
 QDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFI
 KQYGDCLGDIAARDLICAQKFNGLTVLPLLTDemiaQYTSALLAGTITSGWTFGAGAA
 LQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSTASALGKLQDVV
 NQNAQALNTLVKQLSSNFGAISSVLNDILSRLDPPEAEVQIDRLITGRLQSLQTYVTQQLI
 RAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLSFPQSAPHGVVFLHVTYVPA
 QEKNTTAPAICHGDKAHFPREGVVFVSNGTHWFVTQRNFYEPQIITDNTFVSGNCDVV
 IGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVLGDISGINASVVNIQKEIDRLNEV
 AKNLNEIDLQELGKYEYIKWPWYIWLGFIAGLIAIVMTIMLCCMTSCCSCCLKGCC
 SCGSCCKFDEDDSEPVLKGVKLHYT (SEQ ID NO: 4).

[00174] DNA or RNA sequence of S-(deg-RBD) for Delta strain (from 5'-end to 3'-end):

atgttcgtcttctggtcctgctgcctctggtctcctcacagtgcgtcaatctgcgaactcggactcagctgccacctgcttataactaatagcttc
 accagaggcgtgtactatcctgacaaggtgttagaagctcctgctgcactctacacaggatctgtttctgccattcttagcaacgtgacct
 ggttcacgccatccacgtgagcggcaccatggcacaagcggctcgacaatcccgctgctgccttttaacgatggcgtgtacttcgctct
 atcgagaagagcaacatcatcagaggctggatctttggcaccacactggactccaagacacagtctctgctgatcgtgaacaatgccacca
 acgtggtcatcaaggtgtgcgagttccagttttgtaatgatcccttcctggacgtgtactatcacaagaacaataagagctggatggagtccg
 gagtgtattctagcgcacaactgcacatttgagtacgtgagccagcctttcctgatggacctggagggcaagcagggcaattcaagaa
 cctgagggagttcgtgtttaagaatatcgacggctacttcaaaatctactctaagcacacccccatcaacctggcgcgacctgcctcagg
 gcttcagcgccctggagcccctgggtgatctgcctatcggcataacatcacccgggttcagacactgctggccctgcacagaagctacct
 gacacccggcgactcctctagcggatggaccgcccggcgtgccgcctactatgtgggctacctccagccccggaccttctgctgaagta
 caacgagaatggcaccatcacagacgcagtgattgcgcccctggacccccctgagcgagacaaagtgtactgaagtcctttaccgtgga
 gaagggcatctatcagacatccaattcaggggtgcagccagccgaggcgatcgtgcgctttcctgaaatcacaacctgtgccatttggc
 gaggtgttcgaggcaaccgcttcgccagcgtgtacgcctggaataggaagcggatcagcaactgcgtggccgactatagcgtgctgtac

aactccgcctctttcagcacctttaagtgtatggcgtgtccccacaaagctgaatgacctgtgctttaccaacgtctacgccgattcttctgt
gatcagggggcgacgaggtgcgccagatcgccccggccagacaggcaagatcgagactacaattataagctgccagacgatttcacc
ggctgcgtgatcgctggaacagcaacaatctggattccaaagtggcggaactacaattatcggtaccggctgttagaaagagcaatc
tgaagcccttcgagaggacatctctacagaaatctaccaggccggcagcaagccttgcaatggcgtggagggtttaactgttatttcca
ctccagtctacggcttcagcccacaaacggcgtgggctatcagccttaccgctggtggtgctgagctttgagctgtgcacgccccag
caacagtgtgcggccccaagaagtccaccaatctggtaagaacaagtgcgtgaactcaactcaacggcctgaccggcacaggcgtg
ctgaccgagtcacaagaagtctctgcatctcagcagttcggcaggacatcgagataccacagacggcgtgcgcgacccacagac
cctggagatcctggacatcacacctgtcttctggcggtgagcgtgatcacaccggcaccaatacaagcaaccaggtggccgtgct
gtatcagggcgtgaattgtaccgaggtgcccgtggctatccacgccgatcagctgacccaacatggcggtgtacagaccggctccaa
cgtcttcagacaagagccggatgcctgatcggagcagagcacgtgaacaattcctatgagtgcgacatccaatcggcgcggcatctg
tgctcttaccagaccagacaaactctcgagaagagcccgagcgtggcctccagtctatcgcctatacatgtccctggggcgcc
gagaacagcgtggcctacttaacaatagcatcgccatcccaaccaacttcacaatctctgtgaccacagagatcctgccgtgtccatgac
caagacatctgtggactgcacaatgtatatctgtggcgattctaccgagtcagcaacctgtgctccagtacggcagctttgtaccagct
gaatagagccctgacaggcatcgccgtggagcaggataagaacacacagagggtgttcgccagggtgaagcaaatctacaagaccccc
cctatcaaggactttggcggtcctaattttccagatcctgctgatccatccaagccttctaagcgagctttatcgaggacctgtgttaa
caaggtgacctggccgatcgccggtcatcaagcagtatggcgattgctggcgacatcgagccagggacctgatctgcgccaga
agttaatggcctgacctgtgccaccctgctgacagatgagatgatcgacagtacacaagcgccctgctggccggcaccatcacatc
cggtgagaccttggcgagcagggccctccagatcccccttgccatgcagatggcctataggttaacggcatcgccgtgacccagaa
tgtgtgtacgagaaccagaagctgatcgccaatcagtttaactccgcatcggaagatccaggacagcctgtcctctacagccagcgcc
ctgggcaagctccagaatgtgtgaatcagaacgcccaggccctgaatacctggtgaagcagctgagcagcaacttcggcgccatctct
agcgtgtgaatgacatcctgagccggctggacaagggtggaggcagaggtgcagatcgaccggctgatcaccggccggctccagagc
ctccagacctatgtgacacagcagctgatcagggccggcagatcagggccagcgccaatctggcagcaaccaagatgtccgagtgctg
gtgggcccagctaaagagtggtgactttgtggcaagggtatcacctgatgtccttccctcagctgtccccacagcggtggtgtttctgca
cgtgacctacgtccccggcaggagaagaacttcaccacagcccctgccatctgccacgatggcaaggccccactttccaaggaggggc
gtgttcgtgtccaacggcaccactgtgttgacacagcgcaatttctacagccccagatcatcaccacagacaacaccttcgtgagcg
gcaactgtgacgtggtcatcgccatcgtgaacaatacctgtatgatccactccagcccagctggacagctttaaggaggagctggataa
gtatttcaagaatcacacctccccctgacgtggatctggcgacatcagcgcatcaatgcctccgtggtgaacatccagaaggagatcgac
cgctgaacgaggtggctaagaatctgaacgagagcctgatcgacctcaggagctgggcaagtatgagcagatcaatgaatggccctg
gtacatctggctgggcttcatcgccggcctgatcgccatcgtgatggtgacctcatgctgtgtgtatgacatcctgtgttcttgccctgaag
ggctgtgtatgctgtggctcctgtgtaagtttgacgaggatgactctgaacctgtgtgaaggcggtgaagctgcattacacctaa

(SEQ ID NO: 21)

[00175] Protein sequence of S-(deg-RBD) for Delta strain (from N-terminus to C-terminus):

MFVFLVLLPLVSSQCVNLRTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFS
 NVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASIEKSNIIRGWIFGTTLDSKTQSLIV
 NNATNVVIKVCEFQFCNDPFLDVYYHKNNKSWMESGVYSSANNCTFEYVSQPFLMDL
 EGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQT
 LLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSE
 TKCTLKSFTVEKGIYQTSNFRVQPAEAIVRFPQITNLCPFGEVFQATRFASVYAWNKRRI
 SNCVADYSVLVNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVQRQIAPGQTGK
 IADYNYKLPPDDFTGCVIAWNSNNLDSKVGGNYNYRYRLFRKSNLKPFERDISTEIQAG
 SKPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVLSFELLHAPATVCGPKKSTNLV
 KNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGG
 VSVITPGTNTSNQVAVLYQGVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGA
 EHVNNSECDIPIGAGICASYQTQTSNRRRARSVASQSIIAYTMSLGAENSVAYSNNIAI
 PTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLQYGSFCTQLNRALTGIAVEQD
 KNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQ
 YGDCLGDIAARDLCAQKFNGLTVLPPLLDEMIAQYTSALLAGTITSGWTFGAGAALQ
 IPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSTASALGKLQNVVNQ
 NAQALNTLVKQLSSNFGAISSVLNDILSRDLKVEAEVQIDRLITGRLQSLQTYVTQQLIR
 AAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQE
 KNFTTAPAICHGDKAHFPREGVFSNGTHWFVTQRNFYEPQIITDNTFVSGNCDVVIGI
 VNNTVYDPLQPELDSFKEELDKYFKNHTSPDVLDGDISGINASVVNIQKEIDRLNEVAK
 NLNESLIDLQELGKYEYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCLKGCCSC
 GSCKFDEDDSEPVLKGVKLHYT (SEQ ID NO: 22)

[00176] DNA or RNA sequence of S-(deg-RBD) for Wuhan S-2P strain (from 5'-end to 3'-end):

atgttcgtcttcctggctcctgctgcctcctgctcctcacagtgcgtcaatctgacaactcggactcagctgccacgtcttataactaatagcttc
 accagaggcgtgtactatcctgacaaggtgttagaagctccgtgctgcactctacacaggatctgtttctgccattcttagcaacgtgacct
 ggttcacgccatccacgtgagcggcaccatggcacaagcgggtcgacaatcccgtgctgccttttaacgatggcgtgtacttcgcctct
 accgagaagagcaacatcatcagaggctggatcttggcaccacactggactccaagacacagtctctgctgatcgtgaacaatgccacc
 aacgtggtcatcaaggtgctgcgagttccagtttgaatgatccctcctggcgtgtactatcacaagaacaataagagctggatggagtc
 gagtttagagtgtattctagcgccaacaactgcacatttgagtacgtgagccagccttctgatggacctggagggcaagcagggcaattt
 caagaacctgagggagttcgtgtttaagaatatcgacggctacttcaaatctactctaagcacacccccatcaacctgggtgcgcgacctgc
 ctgagggcttcagcgccctggagcccctggatctgcctatcgcatcaacatcacccgggttcagacactgctggccctgcacagaag
 ctacctgacacccggcactcctctagcggatggaccgcccgcgtgccgcctactatgtgggctacctccagccccggaccttctgct
 gaagtacaacgagaatggcaccatcacagacgcagtggtgctgcgcctggacccctgagcgagacaaagtgtactgaagtcctttac

cgtaggaaggcatctatcagacatccaatttcagggtgcagccagccgagcgatcgtagcttctcctgaaatcacaaacctgtgccca
ttggcgagggtgttcgaggcaaccgcttcgccagcgtgtacgcctggaataggaagcggatcagcaactgcgtggccgactatagcgtg
ctgtacaactccgcctctttcagcacctttaagtctatggcgtgtccccacaaagctgaatgacctgtgctttaccaactctacgccgattc
ttcgtgatcagggcgacgaggtgcgccagatgccccggccagacaggcaagatcgagactacaattataagctgccagacgattt
caccggctgcgtgatcgctggaacagcaaatctggattccaagtggcgaggcaactacaattatctgtaccggctgttagaaagagc
aatctgaagcccttcgagaggacatctctacagaaatctaccaggccggcagcaccccttgcaatggcgtggagggtttaactgttattt
cccactccagtcctacggcttcagcccacaaacggcgtgggtatcagccttaccgctgggtgctgagctttgagctgtgcacggc
ccagcaacagtgtagcgcccaagaagtccaccaatctggtgaagaacaagtgctgaactcaactcaacggcctgaccggcacagg
cgtgtgtagccgagtcacaagaagttcctgccatttcagcagttcggcaggacatcgagataccacagacggcgtgcgcgaccac
agacctggagatcctggacatcacacctgctcttcggcggtgagcgtgatcacaccggcaccaatacaagcaaccaggtggcc
gtgctgtatcaggacgtgaattgtaccgaggtgccgtggctatccacggcagatcagctgaccccaacatggcggtgtacagaccggc
tccaacgtcttcagacaagagccggatgcctgatcgagcagagcacgtgaacaattcctatgagtcgacatccaatcggcgcggc
atctgtgctcttaccagaccagacaaactctccagaagagcccgagcgtggcctccagctatcatcgctatacatgtccctggg
cgccgagaacagcgtggcctactctaacaatagcatcgccatcccaaccaacttcacaatctctgtgaccacagagatcctgccgtgtcc
atgaccaagacatctgtggactgcacaatgtatatctgtggcgattctaccgagtgagcaacctgctgtccagtagcgagctttgtacc
cagctgaatagagccctgacaggcatcgccgtggagcaggataagaacacacaggaggtgttcgccaggtgaagcaaatctacaaga
ccccctatcaaggactttggcggtcaattttccagatcctgctgatccatccaagccttctaagcgagctttatcgaggacctgt
gttcaacaaggtgacctggccgatccggcttcacaaagcagtagtgccgattgcctggcgacatcgagccaggacctgatctgcgc
ccagaagttaatggcctgacctgtgccacccctgctgacagatgagatgatcgacagtacacaagcgccctgctggcggccacat
cacatccggatggaccttcggcgaggagccgcccctccagatcccccttgcctatgagatggcctataggttcaacggcatcgccgtgac
ccagaatgtgctgtacgagaaccagaagctgatcgccaatcagtttaactccgcatcggaagatccaggacagcctgtcctctacagcc
agcgccctgggcaagctccaggatgtggtgaatcagaacgcccaggccctgaataacctggtgaagcagctgagcagcaacttcggcg
ccatctctagcgtgctgaatgacatcctgagccggctggaacccggcagaggtgcagatcgaccggctgatcaccggccggct
ccagagcctccagacctatgtgacacagcagctgatcaggccgcccagatcaggccagcgcaatctggcagcaaccaagatgtcc
gagtgctgtgctggccagctaaagagtgtagcttttgggcaagggtatcacctgatgtccttcctcagctgtccccacacggcgtggt
gtttctgcacgtgacctacgtgcccggccaggagaagaacttcaccacagcccctgcatctgcccagatggcaaggcccactttccaag
ggagggcggtgttcgtgtccaacggcaccactgggttgtagacacagcgcaattttacgagccccagatcatcaccacagacaacacctc
gtgagcggcaactgtgacgtggtcatcggtatgaacaataccgtgtatgactccagcccagctggacagctttaaggaggag
ctggataagtatttcaagaatcacacctcccctgacgtggatctggcgacatcagcgcatcaatgcctccgtggtgaacatccagaagg
agatcgaccgctgaacgaggtggctaagaatctgaacgagagcctgatcgacctccaggagctgggcaagtatgagcagtagcatcaag
tggccctgtgacatctggctgggcttcacgcccgtgatcgccatcgtgatggtgacctcatgtgtgtgtatgacatcctgtgttctt
gcctgaagggtgctgtagctgtggctcctgctgaagtttgacgaggatgactctgaacctgtgctgaaggcgctgaagctgcattacacc
taa (SEQ ID NO: 23)

[00177] Protein sequence of S-(deg-RBD) for Wuhan S-2P strain (from N-terminus to C-terminus):

MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFS
 NVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTSEKSNIIRGWIFGTTLDSTQSLIV
 NNATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMD
 LEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQ
 TLLALHRSYLTTPGDSSSGWTAGAAAYVGYLQPRTFLLKYNENGTITDAVDCALDPLS
 ETKCTLKSTVEKGIYQTSNFRVQPAEAIVRFPQITNLCPFGEVFQATRFASVYAWNKRK
 ISNCVADYSVLVNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVQRQIAPGQTG
 KIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNLYRLFRKSNLKPFERDISTEYQA
 GSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVLSFELLHAPATVCGPKKSTNL
 VKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSF
 GVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIG
 AEHVNNSYECDIPIGAGICASYQTQTNSPRRARSVASQSIIAYTMSLGAENSVAYSNN
 SI AIPNTFTISVTTEILPVSMKTSTVDCTMYICGDSTECNLLLQYGSFCTQLNRALTGIAVE
 QDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVT
 LADAGFI KQYGDCLGDIAARDLICAQKFNGLTVLPLLTDemiaQYTSALLAGTITSGWTFGAGAA
 LQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSTASALGKLQDVV
 NQNAQALNTLVKQLSSNFGAISSVLNDILSRDPEAEVQIDRLITGRLQSLQTYVTQQLI
 RAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPA
 QEKNFTTAPAICHGDKAHFPREGVFVSNGTHWFTVQRNFYEPQIITDNTFVSGNCDVV
 IGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVLGDISGINASVNIQKEIDRLNEV
 AKNLNEIDLQELGKYEQYIKWPWYIWLGFIAGLIAIVMTIMLCCMTSCCSCCLKGCC
 SCGSCCKFDEDDSEPVLKGVKLHYT (SEQ ID NO: 24)

[00178] DNA or RNA sequence of S-(deg-RBD) for Delta S-2P strain (from 5'-end to 3'-end):

atgttcgtcttctggtcctgctgcctctggtctcctcacagtgcgtcaatctcgaaactcggactcagctgccacctgcttataactaatagcttc
 accagaggcgtgtactatcctgacaaggtgttagaagctcgtgctgactctacacaggatctgtttctgccattcttagcaacgtgacct
 ggttcacgccatccacgtgagcggcaccatggcacaagcggctcgacaatcccgctgctgccttttaacgatggcgtgtacttcgcctct
 atcgagaagagcaacatcatcagaggctggatctttggcaccacactggactccaagacacagtctctgctgatcgtgaacaatgccacca
 acgtggtcatcaaggtgtgcgagttcagttttgtaatgatcccttcctggacgtgtactatcacaagaacaataagagctggatggagtcgg
 gagtgattctagcggcaacaactgcacatttgagtacgtgagccagccttctctgatggacctggagggaagcagggaatttcaagaa
 cctgaggaggttcgtgtttaagaatatcgacggctacttcaaaatctactctaagcacacccccatcaacctggtgcgacacctgcctcagg
 gcttcagcgccttgagcccctggtggtatctgcctatcgcatcaacatcacccggttcagacactgctggccctgcacagaagctacct

gacacccggcgactcctctagcggatggaccgcccggcgctgccgcctactatgtgggctacctccagccccggaccttctgctgaagta
caacgagaatggcaccatcacagacgcagtggtgattgcgccctggacccccctgagcgcagacaaagtgtactgaagtcctttaccgtgga
gaagggcatctatcagacatccaatttcagggtgcagccagccgaggcgatcgtgcgcttctgaaatcacaacctgtgcccatttggc
gaggtgttcgaggcaaccgcttcgccagcgtgtacgcctggaataggagcgatcagcaactgcgtggccgactatagcgtgctgtac
aactccgctctttcagcaccttaagtgtatggcgtgtccccacaaagctgaatgacctgtgctttaccaacgtctacgccgattcttctgt
gatcagggggcgacgaggtgcgccagatccccccggccagacaggcaagatcgagactacaattataagctgccagacgatttcacc
ggctgcgtgatcgctggaacagcaacaatctggattccaaagtggcggaactacaattatcggtaccggctgttagaagagcaatc
tgaagcccttcgagaggacatctctacagaaatctaccaggccggcagcaagccttgcaatggcgtggagggtttaactgttatttcca
ctccagtcctacggcttcagcccacaaacggcggtgggtatcagccttaccgctggtggtgctgagctttgagctgtgcacgccccag
caacagtgtgcggcccaagaagtccaccaatctggtgaagaacaagtgcgtgaactcaactcaacggcctgaccggcacaggcggtg
ctgaccgagtcacaagaagtctcctgccatttcagcagttcggcaggacatcgagataccacagacgccgtgcgcgaccacagac
cctggagatcctggacatcacacctgctctttcggcggtgagcgtgatcacaccggcaccaatacaagcaaccaggtggccgtgct
gtatcaggcggtgaattgtaccgaggtgcccgtggctatccacgccgatcagctgacccaacatggcggtgtacagaccggctccaa
cgtcttcagacaagagccggatgcctgatcggagcagagcacgtgaacaattcctatgagtgcgacatccaatcggcgccggcatctg
tgcctcttaccagaccagacaaactctcgagaagagcccgagcgtggcctccagcttatcatcgctatacatgtcctggggcgcc
gagaacagcgtggcctacttaacaatagcatcgccatcccaaccaactcacaatctctgtgaccacagagatcctgccggtgctcatgac
caagacatctgtggactgcacaatgtatatctgtggcgattctaccgagtcgagcaacctgctgctccagtagcgagctttgtaccagct
gaatagagccctgacaggcatcgccgtggagcaggataagaacacacaggaggtgttcgcccaggtgaagcaaatctacaagaccccc
cctatcaaggactttggcggttcaattttccagatcctgctgatccatccaagccttcaagcggagctttatcgaggacctgtgttcaa
caaggtgacctggccgatcgccggttcataagcagtatggcgattgcctggcgacatcgagccaggagacctgatctgcgcccaga
agttaatggcctgacctgctgccacccctgctgacagatgagatgatcgacagtacacaagcgccctgctggccggcaccatcacatc
cggatggaccttcggcgagcaggagccgctccagatcccccttgcctatgagatggcctataggttcaacggcatcgccgtgacccagaa
tgtgtgtacgagaaccagaagctgatcgccaatcagtttaactccgcatcggaagatccaggacagcctgtcctctacagccagcgcc
ctgggcaagctccagaatgtggtgaatcagaacgccagggccctgaataacctggtgaagcagctgagcagcaacttcggcgccatctct
agcgtgctgaatgacatcctgagccggctggacccggcagaggtgcagatcgaccggctgatcaccggccggctccagagc
ctccagacctatgtgacacagcagctgatcagggccggcagatcagggccagcgccaatctggcagcaaccaagatgtccgagtgctg
gctgggcccagctaaagagtggtgactttgtggcaagggtatcacctgatgtccttccctcagctgtccccacacggcggtggtgttctgca
cgtgacctacgtgcccggccaggagaagaacttcaccacagcccctgccatctgccacgatggcaaggcccactttcaaggaggggc
gtgttcgtgtccaacggcaccactggtttgtgacacagcgcaattttacgagccccagatcatcaccacagacaacaccttcgtgagcg
gcaactgtgacgtggtcatcgccatcgtaacaataccgtgtatgatccactccagcccagctggacagctttaaggaggagctggataa
gtatttcaagaatcacacctcccctgacgtggatctggcgacatcagcgccatcaatgcctccgtggtgaacatccagaaggagatcgac
cgctgaacgaggtggctaagaatctgaacgagagcctgatcgacctccaggagctgggcaagtatgagcagtagcatcaagtggccctg
gtacatctggctgggcttcacgcccgcctgatcgccatcgtgatggtgacctcatgctgtgctgtatgacatcctgctgttcttgctgaag
ggctgctgtagctgtggtcctgctgtgaagttgacgaggatgactctgaacctgtgctgaaggcgctgaagctgcattacacctaa

(SEQ ID NO: 25)

[00179] Protein sequence of S-(deg-RBD) for Delta S-2P strain (from N-terminus to C-terminus):

MFVFLVLLPLVSSQCVNLRTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFS
 NVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASIEKSNIIRGWIFGTTLDSKTQSLIV
 NNATNVVIKVCEFQFCNDPFLDVYYHKNNKSWMESGVYSSANNCTFEYVSQPFLMDL
 EGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQT
 LLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSE
 TKCTLKSFTVEKGIYQTSNFRVQPAEAIVRFPQITNLCPFGEVFQATRFASVYAWNKRRI
 SNCVADYSVLVNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVQRQIAPGQTGK
 IADYNYKLPPDDFTGCVIAWNSNNLDSKVGGNYNRYRLFRKSNLKPFERDISTEIYQAG
 SKPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVLSFELLHAPATVCGPKKSTNLV
 KNCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGG
 VSVITPGTNTSNQVAVLYQGVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGA
 EHVNNSEYCDIPIGAGICASYQTQTSNRRRRARSVASQSIIAYTMSLGAENSVAYSNNIAI
 PTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLQYGSFCTQLNRALTGIAVEQD
 KNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTADAGFIKQ
 YGDCLGDIAARDLCAQKFNGLTVLPLLTDemiaQYTSALLAGTITSGWTFGAGAALQ
 IPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSTASALGKLQNVVNQ
 NAQALNTLVKQLSSNFGAISSVLNDILSRLDPPAEAEVQIDRLITGRLQSLQTYVTQQLIRA
 AEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEK
 NFTTAPAICHGKAHFPREGVFVSNGTHWFTQRNFYEPQIITDNTFVSGNCDVVIGIV
 NNTVYDPLQPELDSFKEELDKYFKNHTSPDVLGDISGINASVVNIQKEIDRLNEVAKNL
 NESLIDLQELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCCLKGCCSCGSC
 CKFDEDDSEPVLKGVKLHYT (SEQ ID NO: 26)

[00180] DNA or RNA sequence of S-(deg-S2) for Wuhan strain (from 5'-end to 3'-end):

[00181] atgttcgtcttcctggtcctgctgcctctggtctcctcacagtgcgtcaatctgacaactcggactcagctgccacctgc
 ttatactaataagcttcaccagaggcgtgtactatcctgacaaggtgttagaagctccgtgctgcactctacacaggatctgtttctgccattctt
 agcaacgtgacctggtccacgccatccacgtgagcggcacaatggcacaagcgggtcgacaatcccgtgctgccttttaacgatggc
 gtgtacttcgcctctaccgagaagagcaacatcatcagaggctggatctttggcaccacactggactccaagacacagtctctgctgatcgt
 gaacaatgccaccaacgtggtcatcaaggtgtgcgagttccagtttgaatgatcccttcctggcgctgtactatcacaagaacaataagag
 ctggatggagtcgagtttagagtgtattctagcgccaacaactgcacatttgagtacgtgagccagccttctgatggacctggagggca
 agcagggcaatttcaagaacctgagggagttcgtgttaagaatatcgacggctacttcaaaatctacttaagcacacccccatcaacctg
 gtgcgcgacctgcctcagggttcagcgccctggagcccctggtgatctgcctatcggcacatcaacatccccgtttcagacactgctgg

ccctgcacagaagctacctgacacccggcgactcctctagcggatggaccgcccggcgctgccgctactatgtgggtacctccagccc
cggaccttctgtgaagtacaacgagaatggcaccatcacagacgcagtggtgattgcgcctggacccctgagcgagacaaagtgtac
actgaagtcctttaccgtggagaaggcatctatcagacatccaatttcagggtgcagccaaccgagtcctatcgtgcgctttcctaatacac
aaacctgtgcccatttggcgaggtgtcaacgcaacccgcttcgccagcgtgtacgcctggaataggaagcggatcagcaactgcgtggc
cgactatagcgtgctgtacaactcgcctctttcagcaccttaagtctatggcgtgtccccacaaagctgaatgacctgtgctttaccaac
gtctacgccgattcttctgtgatcagggcgacgaggtgcgccagatcgccccggccagacaggcaagatcgagactacaattataag
ctgccagacgatttaccggctgcgtgatcgctggaacagacaactctggattccaaagtggcgaggcaactacaattatctgtaccggct
gtttagaagagcaatctgaagcccttcgagaggacatctctacagaaatctaccaggccggcagcaccccttgaatggcggtggagg
ctttaactgttattccactccagtcctacggcttcagcccacaaacggcggtggctatcagccttaccgctggtggtgctgagctttgag
ctgctgcacgcccagcaacagtgctgcggcccaagaagtccaccaatctggtgaagaacaagtgcgtgaactcaactcaacggcctg
accggcacaggcgtgctgaccgagtcacaagaagttcctgccatttcagcagttcggcaggacatcgagataccacagacgccgt
gcgcgaccacagaccctggagatcctggacatcacacctgctcttccggcggtgagcgtgatcacccggcaccaatacaagca
accaggtggcggtgctgtatcaggacgtgaattgtaccgaggtgccggtgctatccacgccgatcagtgacccaacatggcggtgt
acagcaccggctcaacgtcttcagacaagagccggatgcctgatcggagcagagcacgtgaacaattcctatgagtgcacatccca
atcgcgccggcatctgtgcctttaccagaccagacaaactctccagaagagcccgagcgtggcctccagtcctatcatgcctata
ccatgtccctggcgccgagaaacagcgtggcctactctcagaaatagcatcgccatcccaaccagttcacatctctgtgaccacagagat
cctgcccgtgtccatgaccaagacatctgtggactgcacaatgtatatctgtggcgattctaccgagtgacagcaacctgctgctccagtacg
gcagcttttaccagctgaatagagccctgacaggcatcgccgtggagcaggataagaacacacaggaggtgttcccccaggtgaag
caaatctacaagacccccctatcaaggactttggcggttcacatttccagatcctgcctgatccatcaagccttcaagcggagcttta
tcgaggacctgctgttcaacaaggtgaccttggccgatcgccgcttcaagcagtatggcgattgcctggcgacatcgagccaggg
acctgatctgcgcccagaagttaattggcctgacctgctgccacctgctgacagatgagatgatcgacagtacacaagcgccctgct
ggccggcaccatcacatccggatggaccttggcgagcggccctccagatccctttgccatgcagatggcctataggttcaacgg
catcgcggtgaccagaatgtgctgtacgagaaccagaagctgatcgccaatcagtttaactccgccatcggaagatccaggacgcct
gtcctctacagccagcgcctgggcaagctccaggatgtggtgaatcagaacgccagccctgaataacctgtgaagcagctgagca
gcaacttggcgccatctctagcgtgctgaatgacatcctgagccggctggacaaggtggaggcagaggtgcagatcgaccggctgatc
accggccggctccagagcctccagacctatgtgacacagcagctgatcagggccgagatcagggccagcgccaatctggcagca
accaagatgtccagtgctgtggccagcttaagagagtgacttttggcaaggctatcacctgatgtccttccctcagctgtcccc
acacggcgtggtgttctgcagtgacctacgtgcccggcaggagaagcagttaccacagcccctgccatctgccagatggcaagg
cccactttcaaggaggggcgtgttcgtgtccagggcaccactggtttgtgacacagcgcaatttctacgagccccagatcatcaccac
agacaacaccttcgtgagcggcaactgtgacgtggtcatcgccatcgtgcagaataccgtgtatgatccactccagcccagctggacag
cttaaggaggagctggataagtatttaagcaacacctccctgacgtggatctggcgacatcagcgcatcgaagcctccgtggtg
aacatccagaaggagatcgaccgctgaacgaggtggctaagaatctgcaggagagcctgatcgacctccaggagctgggcaagtatg
agcagtatcaagtggccctgtacatctggctgggtctcatcgccgctgatcccatcgtgatggtgacctcatgctgtgctgtatga

catcctgctgttctgcctgaaggctgctgtagctgtggctcctgctgtaagttgacaggatgactctgaacctgtgctgaaggcgctga
agctgcattacacctaa (SEQ ID NO: 5)

[00182] Protein sequence of S-(deg-S2) for Wuhan strain (from N-terminus to C-terminus):

[00183] MFVFLVLLPLVSSQCVNLTTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHS
TQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTL
DSKTQSLIVNNAITNVVIKVFCEQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTF
EYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVD
LPIGINITRFQTLALHRSYLTGDSGSGWTAGAAAYVGYLQPRFTLLKYNENGTITDA
VDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFAS
VYAWNRRKISNCVADYSVLVNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEV
RQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGNYNYLYRLFRKSNLKPFE
RDISTEIQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVLSFELLHAPATV
CGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQFGRDIADTTDAVRDPQTLE
ILDITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGNSVF
QTRAGCLIGAHEVNNSYECDIPIGAGICASYQTQTNSPRRARSVASQSIIAYTMSLGAEN
SVAYSQNSIAIPTQFTISVTTEILPVSMTKTSVDCTMYICGDSCECSNLLLQYGSFCTQLN
RALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGFQFSQILPDPSKPSKRSFIEDLLFNK
VTLADAGFIKQYGDCLGDIAARDLCAQKFENGLTVLPPLLTDemiaQYTSALLAGTITSG
WTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSTASA
LGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRDPPEAEVQIDRLITGRLQSL
QTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLSFPQSAPHGVVF
LHVTYVPAQEKQFTTAPAICHGDKAHFPREGVFSQGTHWFVTQRNFYEPQIITDNTF
VSGNCDVVIGIVQNTVYDPLQPELDSFKEELDKYFKQHTSPDVLGDISGIQASVVNIQ
KEIDRLNEVAKNLQESLIDLQELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTS
CCSCLKGCCSCGSCCKFDEDDSEPVLKGVKLHYT (SEQ ID NO: 6)

[00184] DNA or RNA sequence of S-(deg-S2) for Delta strain (from 5'-end to 3'-end):

atgttcgttctcctgctgcctctggtctcctcacagtgcgtcaatctgcgaactcggactcagctgccacctgcttataactaatagcttc
accagaggcgtgtactatcctgacaaggtgttagaagctccgtgctgcactctacacaggatctgttctgccattcttagcaacgtgacct
ggttcacgccatccacgtgagcggcaccaatggcacaagcgggtcgacaatcccgtgctgccttttaacgatggcgtgtacttcgcctct
atcgagaagagcaacatcatcagaggctggatcttggcaccacactggactccaagacacagtctctgctgatcgtgaacaatgccacca
acgtggtcatcaaggtgtgcgagttccagtttgaatgatccctcctggacgtgtactatcacaagaacaataagagctggatggagtcgg
gagtgattctagcccaacaactgcacattgagtagctgagccagccttctctgatggacctggaggggaagcagggcaatttcaagaa

cttgaggaggttcgtgtttaagaatatcgacggctacttcaaaatctactctaagcacacccccatcaacctggtgcgcgacctgcctcagg
gcttcagcgccttgagcccctggtgatctgcctatcggcatcaacatcacccggttcagacactgctggccctgcacagaagctacct
gacacccggcgactcctctagcggatggaccgcgctgccgcctactatgtgggctacctccagccccggaccttctgctgaagta
caacgagaatggcaccatcacagacgcagtggtgattgcgccctggacccccctgagcgcagacaaagtgtacactgaagtcctttaccgtgga
gaaggcgatctatcagacatccaatttcagggtgcagccaaccgagtgctatcgtgcgtttcctaatacacaaacctgtgccatttgcgga
ggtgttaacgcaacccgcttcgccagcgtgtacgcctggaataggaagcggatcagcaactgcgtggccgactatagcgtgctgtacaa
ctccgctcttcagcacctttaagtgtatggcgtgtccccacaaagctgaatgacctgtgctttaccaacgtctacgccgattcttctgctga
cagggcgacgaggtgcgccagatcgcggccgagacaggcaagatcgagactacaattataagctgccagacgatttcaccggc
tgctgatcgctggaacagcaacaatctggattccaaagtggcgcgcaactacaattatcggtaccggctgtttagaaagagcaatctga
agcccttcgagaggacatctctacagaaatctaccaggccggcagcaagccttgcaatggcgtggagggtttaactgttattccactc
cagtcctacggcttcagcccacaaacggcgtgggctatcagcctaccgctggtggtgctgagctttgagctgctgcacgccccagca
acagtgctgcggcccaagaagtccaccaatctggtgaagaacaagtgcgtgaactcaactcaacggcctgaccggcacaggcgtgct
gaccgagtcacaagaagtctcgtccatttcagcagttcggcaggacatcgagataccacagacggcgtgcgcgaccacagacc
tgagatcctggacatcacacctgctcttcggcgcgctgagcgtatcacacccggcaccaatacaagcaaccagggtggcgtgctgt
atcagggcgtgaattgtaccgaggtccccgtggctatccacggcatcagctgaccccaacatggcgggtgtacagcaccggctccaac
gtcttcagacaagagccgatgcctgatcggagcagagcacgtgaacaattcctatgagtgcgacatccaatcggcgcggcatctgt
gcctcttaccagaccagacaaactctcgagaagacccggagcgtggcctcccagctatcatcgctatacatgtccctggcgcc
gagaacagcgtggcctactctcagaatagcatcgccatccaacccagttcacaatctctgtgaccacagagatcctgccggtgctcatga
ccaagacatctgtgactgcacaatgtatatctgtggcgattctaccgagtgagcaaacctgctgctccagtaggcagctttgtaccagc
tgaatagacccctgacaggcatcgccgtggagcaggataagaacacacaggaggtgtcgccaggtgaagcaaatctacaagacccc
ccctatcaaggactttggcggtccaattttccagatcctgcctgatccatccaagccttcaagcggagctttatcgaggacctgctgttca
acaagggtgacctggccgatcgccggttcacaaagcagtatggcgattgcctggcgacatcgagccaggacgtgatctgcgccag
aagtttaaggcctgacctgctgccacccctgctgacagatgagatgatcgacagtacacaagcgccctgctggccggcaccatcacat
ccggatggaccttcggcgagcagcggccctccagatcccccttggcatgagatggcctataggttcaacggcatcgcggtgacccaga
atgtgctgtacgagaaccagaagctgatcgccaatcagtttaactccgccatcggaagatccaggacagcctgtcctctacagccagcgc
cctgggcaagctccagaatgtggtgaatcagaacgcccaggccctgaataacctggtgaagcagctgagcagcaacttcggcgccatct
ctagcgtgctgaatgacatcctgagccggctggacaagggtggagcagaggtgcagatcgaccggctgatcaccggcggtccagag
cctccagacctatgtgacacagcagctgatcaggcccgagatcaggccagcgccaatctggcagcaaccaagatgtccgagtg
gtgctgggagcttaagagagtggtgactttgtggcaagggtatcacctgatgtccttcctcagctgccccacacggcggtggtgttctg
cacgtgacctacgtgccccggcaggagaagcagttcaccacagccctgcatctgccacgatggcaaggccactttccaaggagg
gcgtgttcgtgtcccagggcacccactggtttgtgacacagcgcaattttacgagccccagatcatcaccacagacaacacctcgtgag
cggcaactgtgacgtggtcatcggcacgtgcagaataacctgtatgatccactccagcccagctggacagctttaaggaggagctggat
aagtatttcaagcaacacacctcccctgacgtggatctggcgacatcagcggcatccaagcctccgtggtgaacatccagaaggagatc
gaccgcctgaacgaggtggctaagaatctgcaggagacctgatcgacctccaggagctgggcaagtatgagcagtagcatcaagtggcc
ctggtacatctggctgggcttcacgcggcctgatcgccatcgtgatggtgacctcatgctgtgctgtatgacatcctgctgttcttcctga

aggcgctgctgtagctgtggctcctgctgtaagtttgacgaggatgactctgaacctgtgctgaaggcgctgaagctgcattacacctaa
(SEQ ID NO: 27)

[00185] Protein sequence of S-(deg-S2) for Delta strain (from N-terminus to C-terminus):

MFVFLVLLPLVSSQCVNLRTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFS
NVTWFHAIHVS GTNGTKRFDNPVLPFNDGVYFASIEKSNIIRGWIFGTTLD SKTQSL LIV
NNATNVVIKVCEFQFCNDPFLDVYYHKNNKSWMESGVYSSANNCTFEYVSQPFLMDL
EGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQT
LLALHRSYLT PGDSSSGWTAGAAAYVGYLQPRTFLLKYNENGTITDAVDCALDPLSE
TKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRRKRIS
NCVADYSVLYNASFSSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEV RQIAPGQTGKI
ADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYRYRLFRKSNLKPFERDISTEIQAG
SKPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVLSFELLHAPATVCGPKKSTNLV
KNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGG
VSVITPGTNTSNQVAVLYQGVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGA
EHVNNSYECDIPIGAGICASYQTQTSNRRRARSVASQSIAYTMSLGAENSVAYSQNSIAI
PTQFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQD
KNTQEVFAQVKQIYKTPPIKDFGGFQFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQ
YGDCLGDIAARDLICAQKFNGLTVLPPLL TDEMIAQYTSALLAGTITSGWTFGAGAALQ
IPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSTASALGKLQNVVNQ
NAQALNTLVKQLSSNFGAISSVLNDILSRDLKVEAEVQIDRLITGRLQSLQTYVTQQLIR
AAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQE
KQFTTAPAICHGDKAHFPREGV FVSQGTHWFVTQRNFYEPQIITDNTFVSGNCDVVIGI
VQNTVYDPLQPELDSFKEELDKYFKQHTSPDVLDGDISGIQASVVNIQKEIDRLNEVAK
NLQESLIDLQELGKYEYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCCLKGCCSC
GSCCKFDEDDSEPVLKGVKLHYT (SEQ ID NO: 28)

[00186] DNA or RNA sequence of S-(deg-S2) for Wuhan S-2P (from 5'-end to 3'-end):

atgttcgtcttctggtcctgctgcctctggtctcctcacagtgcgtcaatctgacaactcggactcagctgccacgtgcttataactaatagcttc
accagaggcggtgactatcctgacaaggtgttagaagctccgtgctgcactctacacaggatctgtttctgccattcttagcaacgtgacct
ggtccacgccatccacgtgagcggcaccaatggcacaaagcgggtcgacaatcccgtgctgcctttaacgatggcgtgtacttcgcctct
accgagaagagacaatcatcagaggctggatcttggcaccacactggactccaagacacagtctctgctgatcgtaacaatgccacc
aacgtggtcatcaaggtgtgcgagttccagtttgaatgatccctcctggcgctgtactatcacaagaacaataagactggatggagtc
gagtttagagtgtattctagcgccaacaactgcacatttgagtacgtgagccagccttctgatggacctggagggaagcagggaattt

caagaacctgagggagttcgtgtttaagaatatcgacggctacttcaaacttactctaagcacacccccatcaacctggtgcgcgacctgc
ctcagggcttcagcgccctggagcccctgggtgatctgcctatcgcatcaacatcacccggttcagacactgctggccctgcacagaag
ctacctgacacccggcgactcctctagcggatggaccgcccggcgctgccgctactatgtgggctacctccagccccggaccttctgct
gaagtacaacgagaatggcaccatcacagacgcagtggtgattgcgcctggacccccctgagcgagacaaagtgtactgaagtccttac
cgtggagaagggcatctatcagacatccaatttcagggtgcagccaaccgagtgctatcgtgcgcttctaatatcacaacctgtgccatt
tggcgaggtgtcaacgcaaccgcttcgccagcgtgtacgcctggaataggaagcggtacagcaactgcgtggccgactatagcgtgct
gtacaactccgctctttcagcaccttaagtgtatggcgtgtccccacaaagctgaatgacctgtgctttaccaacgtctacgccgattctt
tcgtgatcagggcgacgaggtgcgccagatgcccccggccagacaggcaagatcgagactacaattataagctgccagacgatttc
accggctgcgtgatcgcttgaacagcaacaatctggattccaaagtggcgccgcaactacaattatctgtaccggctgttagaaagagca
atctgaagcccttcgagagggacatctctacagaaatctaccaggccggcagcaccccttgaatggcgtggagggctttaactgttatttc
ccactccagtcctacggcttcagcccacaaacggcgtgggctatcagccttaccgctggtggtgctgagctttgagctgctgcacgcc
cagcaacagtgctgcggcccaagaagtccaccaatctggtgaagaacaagtgcgtgaactcaactcaacggcctgaccggcacaggc
gtgctgaccgagtcacaagaagtctcctgcatcttcagcagttcggcaggacatcgagataccacagacggcgtgcgcgaccaca
gacctggagatcctggacatcacacctgctctttcggcggtgagcgtgatcacaccggcaccaatacaagcaaccaggtggccgt
gctgtatcaggacgtgaattgtaccgaggtgcccgtggctatccacgccgatcagtgacccaacatggcggtgtacagaccggctc
caacgtcttcagacaagagccggatgctgatcggagcagagcacgtgaacaattcctatgagtgacatcccaatcgcgccggcat
ctgtgctcttaccagaccagacaaacttcccagaagagccggagcgtggcctccagtcctatcatgcctataacctgtccctgggc
gccgagaacagcgtggcctactctcagaatagcatcgccatcccaaccagttcacaatctctgtgaccacagagatcctgccgtgtcca
tgaccaagacatctgtgactgcacaatgtatatctgtggcgattctaccgagtgacgaacctgctgctccagtacggcagctttgtacc
agctgaatagagccctgacaggcatcgccgtggagcaggataagaacacacaggaggtgttcgccaggtgaagcaatctacaagac
ccccctatcaaggactttggcggttccaattttccagatcctgctgatccatcaagccttctaaggcgagctttatcgaggacctgctg
ttcaacaagggtgacctggccgatgccggcttcatcaagcagtatggcgattgctggcgacatcgagccaggacactgatctgcgc
cagaagttaatggcctgacctgctgccacccctgctgacagatgagatgatcgacagtacacaagcgccctgctggccggcaccatc
acatccggatggaccttcggcgaggagccgctccagatccccttggcatgagatggcctataggttaacggcatcgcgctgacc
cagaatgtgctgtacgagaaccagaagctgatcgccaatcagtttaactccgcatcggaagatccaggacagcctgtcctctacagcca
gcgcccctgggcaagctccaggtgtggtgaatcagaacgcccaggccctgaataacctggtgaagcagctgagcagcaacttcggcgc
catctctagcgtgctgaatgacatcctgagccggctggacccggcgaggcagaggtgcagatcgaccggctgatcaccggccggctc
cagagcctccagacctatgtgacacagcagctgatcaggggccggagatcaggggccagcgccaatctggcagcaaccaagatgtccg
agtgcgtgctggccagcttaagagagtgactttgtggcaagggtatcacctgatgtccttccctcagctgccccacacggcggtgtg
ttctgcacgtgacctagctgccccccaggagaagcagttcaccacagcccctgccatctgccacgatggcaaggcccactttccaagg
gagggcggtgttctgtcccagggcacccactggttgtgacacagcgcaatttctacgagccccagatcatcaccacagacaacaccttcg
tgagcggcaactgtgacgtggtcatcggtatgcagaataccgtgtatgatccactccagcccagctggacagctttaaggaggagct
ggataagtatttcaagcaacacacctcccctgacgtggatctggcgacatcagcgccatccaagcctccgtggtgaacatccagaagga
gatcgaccgctgaacgaggtggctaagaatctgcaggagacgtgatcgacctcaggagctgggcaagtatgacagtatcaagt
ggccctggtacatctgctgggcttcacggccgctgatcgccatcgatggtgacctatgctgtgctgtatgacatcctgctgttcttg

cctgaagggctgctgtagctgtggctcctgctgtaagtttgacgaggatgactctgaacctgtgctgaagggcgtgaagctgcattacacct
aa (SEQ ID NO: 29)

[00187] Protein sequence of S-(deg-S2) for Wuhan S-2P (from N-terminus to C-terminus):

MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFS
NVTWFHAIHVS GTNGTKRFDNPVLPFNDGVYFASTKSNIRGWIFGTTLD SKTQSLIV
NNATNVVIKVCEFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMD
LEGKQGNFKNLREFVFKNIDGYFKIYKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQ
TLLALHRSYLT PGDSSSGWTAGAAAYVGYLQPRTFLLKY NENGTITDAVDCALDPLS
ETKCTLKSF TVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRRKRI
SNCVADYSVL YNSASFSTFKCYGVSP TKLNDLCFTNVYADSFVIRGDEV RQIAPGQTGK
IADYNYKL PDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIQAG
STPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVLSFELLHAPATVCGPKKSTNLV
KNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSF GG
VSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGA
EHVNNSYECDIPIGAGICASYQTQTN SPRRARSVASQSIIAYTMSLGAENSVAYSQNSIAI
PTQFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQD
KNTQEVFAQVKQIYKTPPIKDFGGFQFSQILPDPSKPSKRSFIEDLLFNKVT LADAGFIKQ
YGDCLGDIAARDLICAQKFNGLTVLPPLL TDEMIAQYTSALLAGTITSGWTFGAGAALQ
IPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSTASALGKLQDVVNQ
NAQALNTLVKQLSSNFGAISSVLNDILSRLD PPEAEVQIDRLITGRLQSLQTYVTQQLIRA
AEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEK
QFTTAPAICHGDKAHFPREGVFVSQGTHWFTVQRNFYEPQIITDNTFVSGNCDVVIGIV
QNTVYDPLQPELDSFKEELDKYFKQHTSPD VDLGDISGIQASVVNIQKEIDRLNEVAKNL
QESLIDLQELGKYEQYIKWPWYIWLGFIAGLIAIVMTIMLCCMTSCC SCLKGCCSCGSC
CKFDEDDSEPV LKGVKLHYT (SEQ ID NO: 30)

[00188] DNA or RNA sequence of S-(deg-S2) for Delta S-2P strain (from 5'-end to 3'-end):

atgttcgtcttctggtcctgctgcctctggtctcctcacagtgcgtcaatctcgaaactcggactcagctgccacctgcttataactaatagcttc
accagaggcgtgtactatcctgacaaggtgttagaagctccgtgctgcactctacacaggatctgttctgccattcttagcaacgtgacct
ggttcacgccatccacgtgagcggcaccatggcacaagcgggtcgacaatcccgctgctgccttttaacgatggcgtgtacttcgctct
atcgagaagagcaacatcatcagaggctgcatctttggcaccacactggactccaagacacagtctctgctgatcgtgaacaatgccacca
acgtggtcatcaaggtgtgcgagttccagttttgtaatgatccctcctggacgtgtactatcacaagaacaataagagctggatggagtcgg

gagtgtattctagcgccaacaactgcacatttgagtacgtgagccagccttctgatggacctggagggaagcagggaatttcaagaa
cctgaggggagttcgtgtttaagaatatcgacggctacttcaaatctactctaagcacacccccatcaacctgggtgcgcgacctgcctcagg
gcttcagcgccctggagcccctgggtgatctgcctatcggcataacatcacccgggttcagacactgctggccctgcacagaagctacct
gacacccggcgactcctctagcggatggaccggcgctgccgcctactatgtgggctacctccagccccggaccttctgctgaagta
caacgagaatggcaccatcacagacgcagtggtgattgcgcctggaccccctgagcgcagacaaaagtacactgaagtctttaccgtgga
gaagggcattctatcagacatccaatttcagggtgcagccaaccgagctctatcgtgcgcttcttaataacacaaacctgtgccatttggcga
gggtgttaacgcaacccgcttcgccagcgtgtacgcctggaataggaagcggatcagcaactgcgtggccgactatagcgtgctgtacaa
ctccgctcttccagcacctttaagtgtatggcgtgtccccacaaagctgaatgacctgtgctttaccaacgtctacgccgattcttctgctgat
cagggggcgacgaggtgcgcgacatgccccggccagacaggcaagatcgagactacaattataagctgcagacagatttcaccggc
tgcgtgatcgctggaacagcaacaatctggattccaagtggcgcgcaactacaattatcggtaccggctgtttagaaagagcaatctga
agcccttcgagaggacatctctacagaaatctaccaggccggcagcaagccttgcaatggcgtggagggtttaactgttatttccactc
cagtcctacggcttcagcccacaaacggcgtgggctatcagccttaccgctgggtgctgagctttgagctgctgcacgccccagca
acagtgtgcgcccccaagaagtccaccaatctgtgaagaacaagtgcgtgaactcaactcaacggcctgaccggcacaggcgtgct
gaccgagtcacaagaagtctcctgcatcttcagcagttcggcaggacatcgagataccacagacgctgctgcgacccacagaccc
tgagatcctggacatcacacctgctcttccggcgctgagcgtgatcacacccggcaccaatacaagcaaccagggtggcgtgctgt
atcaggcgctgaattgtaccgaggtccccgtggctatccacggcatcagctgacccaacatggcgggtgtacagcaccggctccaac
gtcttcagacaagagccgatgctgatcgagcagagcagctgaacaattctatagtgcgacatccaatcggcgcggcgtctgt
gcctcttaccagaccagacaaactctcgagaagagcccgagcgtggcctccagctctatcatgcctataccatgtccctggcgcc
gagaacagcgtggcctactctcagaatagcatcgccatcccaacccagttcacaatctctgtgaccacagagatcctgcccgtgtccatga
ccaagacatctgtgactgcacaatgtatatctgtggcattctaccgagtgagcaacctgctgctccagtagggcagcttttaccagc
tgaatagacccctgacaggcatcgccgtggagcaggataagaacacacaggaggtgtcgccagggtgaagcaaatctacaagcccc
ccctatcaaggactttggcggttccaatttccagatcctgcctgatccatccaagccttctaagcggagctttatcaggacctgctgttca
acaagggtgacctggccgatcgccgctcatcaagcagtatggcgattgctggcgacatcgagccaggacgtatctgcgccag
aagtttaattggcctgacctgctgccacccctgctgacagatgagatgatgcacagtacacaagcgccctgctggccggcaccatcacat
ccggatggaccttcggcgaggagccgcccctccagatcccccttggcatgcagatggcctataggttcaacggcatcgcgctgaccaga
atgtgctgtacgagaaccagaagctgatcgcaatcagtttaactccgccatcggaagatccaggacagcctgtcctctacagccagcgc
cctgggcaagctccagaatgtgtgaatcagaacgcccaggccctgaatacctggtgaagcagctgagcagcaacttcggcgccatct
ctagcgtgctgaatgacatcctgagccggctgga^{ccgccc}ggaggcagaggtgcagatcgacggctgatcaccggccggctccaga
gcctccagacctatgtgacacagcagctgatcagggccggcagatcagggccagcgccaatctggcagcaaccaagatgtccgagtg
cgtgctgggcccagctaaagagtggtgactttgtggcaagggtatcacctgatgtccttccctcagctgccccacacggcggtgtgttct
gcacgtgacctacgtgcccggcaggagaagcagttcaccacagcccctgcatctgccacgatggcaaggcccactttcaaggagag
ggcgtgttctgtcccagggcacccactggtttgtgacacagcgcaatttctacagccccagatcatcaccacagacaacaccttctga
gcggcaactgtgacgtggtcatcgccatcgtgcagaataccgtgtatgatccactccagcccagctggacagctttaaggaggagctgg
ataagtatttcaagcaacacacctcccctgacgtggatctggcgacatcagcgcatccaagcctccgtggtgaacatccagaaggagat
cgaccgcctgaacgaggtggctaagaatctgcaggagagcctgatcgacctcaggagctgggcaagtatgagcagtagatcaagtggc

cctggatcatctggctgggcttcacgcgccgctgatgccatcgatggatgaccatcatgctgtgctgtatgacatcctgctgttcttgctg
aagggtgctgtagctgtggctcctgctgtaagttgacgaggatgactctgaacctgtgctgaaggcgctgaagctgcattacacctaa
(SEQ ID NO: 31)

[00189] Protein sequence of S-(deg-S2) for Delta S-2P strain (from N-terminus to C-terminus):

MFVFLVLLPLVSSQCVNLRTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFS
NVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASIEKSNIIRGWIFGTTLDSTQSLIV
NNATNVVIKVCEFQFCNDPFLDVYYHKNNKSWMESGVYSSANNCTFEYVSQPFLMDL
EGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQT
LLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSE
TKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNKRKIS
NCVADYSVLVNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVQRQIAPGQTGKI
ADYNYKLPDDFTGCVIAWNSNNLDSKVGNNYNYRYRLFRKSNLKPFERDISTEYQAG
SKPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVLSFELLHAPATVCGPKKSTNLV
KNKCVNFNFNGLTGTGVLTESNKKFLPFQFGRDIADTTDAVRDPQTLEILDITPCSFGG
VSVITPGTNTSNQVAVLYQGVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGA
EHVNNSYECDIPIGAGICASYQTQTSNRRRARSVASQSIAYTMSLGAENSVAYSQNSIAI
PTQFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQD
KNTQEVFAQVKQIYKTPPIKDFGGFQFSQILPDPSPKRSFIEDLLFNKVTLADAGFIKQ
YGDCLGDIAARDLCAQKFNGLTVLPPLLTDemiaQYTSALLAGTITSGWTFGAGAALQ
IPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSTASALGKLQNVVNQ
NAQALNTLVKQLSSNFGAISSVLNDILSRLDPEAEVQIDRLITGRLQSLQTYVTQQLIRA
AEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEK
QFTTAPAICHGKAHFPREGVFVSQGTHWFTQRNFYEPQIITDNTFVSGNCDVVIGIV
QNTVYDPLQPELDSFKEELDKYFKQHTSPDVLGDISGIQASVVNIQKEIDRLNEVAKNL
QESLIDLQELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCCLKGCCSCGSC
CKFDEDDSEPVLKGVKLHYT (SEQ ID NO: 32)

[00190] DNA or RNA sequence of S-(S2-1194) for Wuhan strain (from 5'-end to 3'-end):

atgttcgtcttcctggctcctgctgcctctggtctcctcacagtgcgtcaatctgacaactcggactcagctgccacctgcttataactaatagcttc
accagaggcgtgtactatcctgacaagggttttagaagctccgtgctgcactctacacaggatctgtttctgccattctttagcaacgtgacct
ggttccacgccatccacgtgagcggcacaatggcacaagcggttcgacaatcccgtgctgccttttaacgatggcgtgtacttcgctct
accgagaagagcaacatcatcagaggctggatctttggcaccacactggactccaagacacagtctctgctgatcgtaacaatgccacc

aacgtggtcatcaaggtgtgaggtccagttttgtaatatgcccttctgggctgtactatcacagaacaataagagctggatggagtcc
gagtttagagtgtattctagcgccaactgcacatttgagtacgtgagccagccttctgatggacctggagggaagcagggaattt
caagaacctgagggagttcgtgttaagaatatcgacggctacttcaaatctacttaagcacacccccatcaacctggtgcgcgacctgc
ctcagggcttcagcgccctggagccccgtggatctgcctatcgcatcaacatcacccggttcagacactgctggccctgcacagaag
ctacctgacacccggcgactcctctagcggtgagccgcccggcgctgccgctactatgtgggctacctccagccccggaccttctgct
gaagtacaacgagaatggcaccatcacagacgcagtgattgcgcctggacccctgagcgagacaaagtgtacactgaagtccttac
cgtggagaaggcatctatcagacatccaatttcagggtgcagccaaccgagctctatcgtgcgcttctcaatatcacaacctgtgccatt
tggcgaggtgttaacgcaaccgcttcgacgctgtacgcctggaataggaagcggatcagcaactgcgtggccgactatagcgtgct
gtacaactccgctcttctcagcaccttaagtgtatggcgtgtccccacaaagctgaatgacctgtgtttaccaacgtctacgccgattctt
tcgtgatcagggcgacgaggtgcgccagatcgccccggccagacaggcaagatcgagactacaattataagctgccagacgatttc
accggctgcgtgatcgcttgaacagcaacaatctgattccaaagtggcgccgaactacaattatctgtaccggctgttagaaagagca
atctgaagcccttcgagaggacatctctacagaaatctaccaggccggcagcaccccttgaatggcgtggagggttttaactgttatttc
ccactccagtcctacggcttcagcccacaaacggcgtgggctatcagccttaccgctggtggtgctgagctttgagctgtgcacgcc
cagcaacagtgtgcggcccaagaagtccaccaatctggtgaagaacaagtgcgtgaactcaactcaacggcctgaccggcacaggc
gtgctgaccgagtcacaagaagtctcctgcatcttcagcagttcggcaggacatcgagataccacagacggctgcgcgaccaca
gacctggagatcctggacatcacacctgctcttccggcgctgagcgtgatcacaccggcaccaatacaagcaaccaggtggccgt
gctgtatcaggacgtgaattgtaccgaggtgcccgtggctatccacgcccagatcagctgacccaacatggcggtgtacagaccggctc
caacgtcttcagacaagagccggatgcctgatcgagcagagcacgtgaacaattcctatgagtgcgacatccaatcgcgccggcat
ctgtgctcttaccagaccagacaaactctccagaagagccggagcgtggcctccagctctatcatcgctataacctgtccctgggc
gccgagaacagcgtggcctactctcagaaatagcatcgccatcccaaccagttcacaatctctgtgaccacagagatcctgccgtgtcca
tgaccaagacatctgtggactgcacaatgtatatctgtggcgattctaccgagtgacgaacctgctgctccagtacggcagctttgtacc
agctgaatagacccctgacaggcatcgccgtggagcaggataagaacacacaggaggtgttcgcccaggtgaagcaaatctacaagac
ccccctatcaaggactttggcgcttcctcatttccagatcctgcctgatccatccaagccttctaagcggagctttatcaggacctgctg
ttcaacaagggtgacctggccgatgcggcttcatcaagcagatgtggcgattgcctggcgacatcgagccaggacactgatctgcgc
cagaagttaatggcctgacctgctgccacccctgctgacagatgagatgatcgacagtacacaagcgccctgctggccggcaccatc
acatccggatggaccttcggcgaggagccgcccctccagatccccttgcctatgcagatggcctataggttaacggcatcgcgctgacc
cagaatgtgctgtacgagaaccagaagctgatcgcaatcagtttaactccgccatcggaagatccaggacagcctgtcctctacagcca
gcgcctgggcaagctccaggatgtggtgaatcagaacgccaggccctgaataccctggtgaagcagctgagcagcaacttcggcgc
catctctagcgtgctgaatgacatcctgagccggctggacaaggtggaggcagaggtgcagatcaccggctgatcaccggccggctcc
agagcctccagacctatgtgacacagcagctgatcagggccgcccagatcagggccagcgccaatctggcagcaaccaagatgtccga
gtgctgctgggcccagtctaagagagtggacttttggcaagggtatcacctgatgtccttcctcagctgtccccacacggcggtggtgtt
tctgcagtgacctacgtgcccggcaggagaagcagttcaccacagcccctgcatctgccacgatggcaaggcccactttccaaggg
agggcggtgtcgtgtccaggggcaccactggtttgtgacacagcgcaatttctacgagccccagatcatcaccacagacaacaccttctg
gagcggaactgtgacgtggtcatcgcatcgtgagaaataccgtgtatgatccactccagcccagctggacagctttaaggaggagct

ggataagtatttcaagcaacacacctcccctgacgtggatctggcgacatcagcggcatccaagcctccgtggtgaacatccagaagga
 gatcgaccgcctgaacgaggtggctaagaatctgaacgagagcctgatcgacctccaggagctgggcaagtatgagcagtacatcaagt
 ggccctggtacatctggctgggtctcatcgccggcctgatcgccatctgatggtgacctcatgctgtgctgtatgacatcctgctgttcttg
 cctgaagggctgctgtagctgtggctcctgctgtaagttgacgaggatgactctgaacctgtgctgaagggcgtgaagctgcattacacct
 aa (SEQ ID NO: 7)

[00191] Protein sequence of S-(S2-1194) for Wuhan strain (from N-terminus to C-terminus):

MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFS
 NVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTKSNIRGWIFGTTLDSTKTSQSLIV
 NNATNVVIKVCFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMD
 LEGKQGNFKNLREFVFKNIDGYFKIYKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQ
 TLLALHRSYLTPGDSSSGWTAGAAAYVGYLQPRTFLLKYNENGTITDAVDCALDPLS
 ETKCTLKSTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNKRKI
 SNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRIAPGQTGK
 IADYNYKLPPDDFTGCVIAWNSNNLDSKVGGNYNLYRLFRKSNLKPFERDISTEIQAG
 STPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVLSFELLHAPATVCGPKKSTNLV
 KNKCVNFNFNGLTGTGVLTESNKKFLPFQFGRDIADTTDAVRDPQTLEILDITPCSFGG
 VSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGA
 EHVNNSEYCDIPGAGICASYQTQTNPPRRARSVASQSIIAYTMSLGAENSVAYSQNSIAI
 PTQFTISVTTEILPVSMKTSTVDCTMYICGDSTECNLLLQYGSFCTQLNRALTGIAVEQ
 DKNTQEVFAQVKQIYKTPPIKDFGGFQFSQILPDPSKPSKRSFIEDLLFNKVTADAGFIK
 QYGDCLGDIAARDLCAQKFNGLTVLPLLTDEMIAQYTSALLAGTITSGWTFGAGAAL
 QIPFAMQMA YRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSTASALGKLQDVVN
 QNAQALNTLVKQLSSNFGAISSVLNDILSRDPPEAEVQIDRLITGRLQSLQTYVTQQLIR
 AAERASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQE
 KQFTTAPAICHGDKAHFPREGVFSQGTHWFTVQRNFIYEPQIITDNTFVSGNCDVVIGI
 VQNTVYDPLQPELDSFKEELDKYFKQHTSPDVLGDISGQASVVNIQKEIDRLNEVAK
 NLNESLIDLQELGKYEYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCLKGCCCCG
 SCCKFDEDDSEPVKGVKLHYT (SEQ ID NO: 8)

[00192] DNA or RNA sequence of S-(S2-1194) for Wuhan S-2P strain (from 5'-end to 3'-end):

atgttcgtcttctggtcctgctgcctctggtctcctcacagtgcgtcaatctgacaactcggactcagctgccacctgcttataactaatagcttc
 accagaggcgtgtactatcctgacaaggtgttagaagctccgtgctgcactctacacaggatctgtttctgccattctttagcaacgtgacct

gggtccacgccatccacgtgagcggcaccaatggcacaagcgggtcgacaatcccgtgctgcctttaacgatggcgtgtacttcgcctct
accgagaagagcaacatcatcagaggctggatcttggcaccacactggactccaagacacagtctctgctgatcgtgaacaatgccacc
aacgtggatcatcaaggtgtgcgagttccagttttgaatgatcccttctgggcgtgtactatcacaagaacaataagagctggatggagtc
gagtttagagtgtattctagcgccaacaactgcacatttgagtacgtgagccagccttctgatggacctggagggaagcagggaattt
caagaacctgagggagtgctgtttaagaatatcgacggctacttcaaatctactctaagcacacccccatcaacctggcgcgacctgc
ctcagggcttcagcgccctggagccccctggatctgcctatcgccatcaacatcacccgggttcagacactgctggccctgcacagaag
ctacctgacacccggcgactccttagcggatggaccgcccggcgtgccctactatgtgggctacctccagccccggaccttctgct
gaagtacaacgagaatggcaccatcacagacgcagtgattgcgccctggacccctgagcgagacaaagtgtacactgaagtccttac
cgtggagaaggcatctatcagacatccaatttcagggtgcagccaaccgagtgctatcgtgcgcttctcaatatcacaacctgtgccatt
tggcgaggtgttaacgcaaccgcttcgccagcgtgtacgcctggaataggaagcggatcagcaactgcgtggccgactatagcgtgct
gtacaactccgctcttccagacctttaagtgtatggcgtgtccccacaaagctgaatgacctgtgctttaccaacgtctacgccgattctt
tcgtgatcagggcgacgaggtgcgccagatcgccccggccagacaggcaagatcgagactacaattataagctgccagacgatttc
accggctgcgtgatcgccctggaacagcaacaatctgattccaaagtggcgccgaactacaattatctgtaccggctgttagaaagagca
atctgaagcccttcgagaggacatctctacagaaatctaccaggccggcagcacccttgcaatggcgtggagggtttaactgttatttc
ccactccagtcctacggcttcagcccacaaacggcgtgggctatcagccttaccgctgggtggtgctgagctttgagctgtgcacggcc
cagcaacagtgctgcggcccaagaagtccaccaatctggtgaagaacaagtgcgtgaactcaactcaacggcctgaccggcacaggc
gtgctgaccgagtcacaagaagtctctccatttcagcagttcggcaggacatcgagataccacagacggctgcgcgaccacaca
gacctggagatcctggacatcacacctgctcttccggcggtgagcgtgatcacccggcaccaatacaagcaaccaggtggccgt
gctgtatcaggacgtgaattgtaccgaggtgcccgtggctatccacggcagatcagtgacccaacatggcgggtgtacagcaccggctc
caacgtcttcagacaagagccggatgcctgatcgagcagagcacgtgaacaattcctatgagtgcgacatcccaatcgccgcccgc
ctgtgctcttaccagaccagacaaactctccagaagagccggagcgtggcctcccagtcctatcatcgctataacctgtccctgggc
ggcgagaacagcgtggcctactctcagaatgcatcgccatcccaaccagttcacaatctctgtgaccacagagatcctggcgtgtcca
tgaccaagacatctgtggactgcacaatgtatatctgtggcgattctaccgagtgacgaacctgctgctccagtacggcagctttgtacc
agctgaatagagccctgacaggcatcgccgtggagcaggataagaacacacaggaggtgttcgcccaggtgaagcaaatctacaagac
ccccctatcaaggactttggcggcttccaattttccagatcctgcctgatccatcaagccttctaagcggagctttatcaggacctgctg
ttcaacaagggtgacctggccgatgccggcttcatcaagcagtatggcgattgcctggcgacatcgacggcaggacctgatctgcgcc
cagaagttaattggcctgaccgtgctgccacccctgctgacagatgagatgatcgacagtacacaagcgccctgctggccggcaccatc
acatccggatggaccttcggcgaggagccgcccctccagatccctttgccatgacagatggcctataggttcaacggcatcgccgtgacc
cagaatgtgctgtacgagaaccagaagctgatcgccaatcagtttaactccgcatcggaagatccaggacagcctgtcctctacagcca
gcgccttgggcaagctccaggatgtggtgaatcagaacgcccaggccctgaataccctggtgaagcagctgagcagcaatttcggcgc
catctctagcgtgctgaatgacatcctgagccggctggacccggcgaggcagaggtgcagatcgacccggctgatcaccggccggctc
cagagcctccagacctatgtgacacagcagctgatcaggccggccgagatcaggccagcgccaatctggcagcaaccaagatgtccg
agtgcgtgctggccagcttaagagagtgacttttggcaagggtatcacctgatgtccttccctcagctgtccccacacggcgtggtg
tttctgcagctgacctagtgccccccaggagaagcagttcaccacagcccctgccatctgccagatggcaaggcccactttccaagg
gagggcgtgttcgtgccaggccaccactggtttgtgacacagcgcaatttctacgagccccagatcatcaccacagacaacaccttcg

tgagcggcaactgtgacgtgggtcatcggtcgcagaataccgtgtatgatccactccagcccagctggacagctttaaggaggagct
 ggataagtatttcaagcaacacacctcccctgacgtggatctggcgacatcagcggtcatccaagcctccgtggtgaacatccagaagga
 gatcgaccgcctgaacgaggtggctaagaatctgaacgagagcctgatcgacctcaggagctgggcaagtatgagcagtagcatcaagt
 ggccctggtacatctggctgggtctcatcgccggcctgatcgccatcgtgatggtgacctcatgctgtgctgtatgacatcctgctgttcttg
 cctgaagggctgctgtagctgtggctcctgctgtaagttgacgaggatgactctgaacctgtgctgaagggcgtgaagctgcattacacct
 aa (SEQ ID NO: 33)

[00193] Protein sequence of S-(S2-1194) for Wuhan S-2P strain (from N-terminus to C-terminus):

MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFS
 NVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTKSNIRGWIFGTTLDSTQSLIV
 NNATNVVIKVFCEFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFMD
 LEGKQGNFKNLREFVFKNIDGYFKIYKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQ
 TLLALHRSYLTTPGDSSSGWTAGAAAYVGYLQPRFTLLKYNENGTITDAVDCALDPLS
 ETKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRRKI
 SNCVADYSVLVNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVQRAPGQTGK
 IADYNYKLPPDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIQAG
 STPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVLSFELLHAPATVCGPKKSTNLV
 KNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGG
 VSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGA
 EHVNNSECDIPIGAGICASYQTQTNSPRRARSVASQSIIAYTMSLGAENSVAYSQNSIAI
 PTQFTISVTTEILPVSMTKTSVDCTMYICGDSTECNLLLQYGSFCTQLNRALTGIAVEQD
 KNTQEVFAQVKQIYKTPPIKDFGGFQFSQILPDPSKPSKRSFIEDLLFNKVTADAGFIKQ
 YGDCLGDIAARDLCAQKFNGLTVLPPLLDEMIAQYTSALLAGTITSGWTFGAGAALQ
 IPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSTASALGKLQDVVNQ
 NAQALNTLVKQLSSNFGAISSVLNDILSRDPPPEAEVQIDRLITGRLQSLQTYVTQQLIRA
 AEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEK
 QFTTAPAICHGDKAHFPREGVFVSQGTHWFTVQRNFYEPQIITDNTFVSGNCDVVIGIV
 QNTVYDPLQPELDSFKEELDKYFKQHTSPDVLGDISGIQASVVNIQKEIDRLNEVAKNL
 NESLIDLQELGKYEYIKWPWYIWLGFIAGLIAIVMTIMLCCMTSCCSCCLKGCCSCGSC
 CKFDEDDSEPVKGVKLHYT (SEQ ID NO: 34)

[00194] DNA or RNA sequence of S-(deg-RBD-801) for Wuhan strain (from 5'-end to 3'-end):

atgttcgtcttcttggtcctgctgcctctggtctcctcacagtgcgtcaatctgacaactcggactcagctgccacctgcttataactaatagcttc
accagaggcgtgtactatcctgacaaggtgttagaagctccgtgctgcactctacacaggatctgtttctgccattcttagcaacgtgacct
ggttcacgccatccacgtgagcggcaccaatggcacaagcggctcgacaatcccgtgctgccttttaacgatggcgtgtacttcgcctct
accgagaagagcaacatcatcagaggctggatcttggcaccacactggactccaagacacagtctctgctgatcgtgaacaatgccacc
aacgtggtcatcaaggtgtgcgagttccagttttgtaatatgatcccttctggcgctgtactatcacaagaacaataagagctggatggagtc
gagtttagagtgtattctagcgccaacaactgcacatttgagtacgtgagccagccttctgatggacctggagggaagcagggaattt
caagaacctgagggagtgctgtttaagaatatcgacggctacttcaaatctacttaagcacacccccatcaacctggtgcgcgacctgc
ctcagggcttcagcgccctggagcccctgggtgatctgcctatcgccatcaacatcacccggttcagacactgctggccctgcacagaag
ctacctgacacccggcgactcctctagcggatggaccgcccggcgctgccctactatgtgggctacctccagccccggaccttctgct
gaagtacaacgagaatggcaccatcacagacgcagtgattgcgcctggacccccctgagcgagacaaagtgtactgaagtcctttac
cgtggagaagggcatctatcagacatccaatttcaggggtgcagcca[gccgag]gctatcgtgcgttttctgaaatcacaaacctgtgcc
atttggcgaggtgttc[gag]gcaacccgcttcgcccagcgtgtacgcctggaataggaagcggatcagcaactgcgtggccgactatagcgt
gctgtacaactccgcctcttfcagcacctttaagtctatggcgtgtccccacaaagctgaatgacctgtgctttaccaacgtctacgccgat
tcttctgctgatcagggcgacgaggtgcgccagatcgccccggccagacaggcaagatcgagactacaattataagctgccagacga
tttaccggctgctgatcgctggaacagcaacaatctggattccaaagtggcgcgcaactacaattatctgtaccggctgtttagaaga
gcaatctgaagcccttcgagaggacatctctacagaatctaccaggccggcagcaccccttgcaatggcgtggagggttttaactgtta
tttccactccagtcctacggcttcagcccacaaacggcggtggctatcagcctaccgcgtgggtggtgctgagctttgagctgctgcacg
ccccagcaacagtgtgcggccccagaagtcaccaatctggtgaagaacaagtgcgtgaacttcaacttcaacggcctgaccggcaca
ggcgtgctgaccgagtcacaagaagtctcctgccatttcagcagttcggcaggacatcgagataccacagacccgtgcgcgaccc
acagacctggagatcctggacatcacacctgctcttctggcgcgctgagcgtgatcacacccggcaccataaagcaaccaggtgg
ccgtgctgtatcaggacgtgaattgtaccgaggtgcccggtgctatccacgccgatcagctgacccaacatggcggtgtacagcacccg
gtccaacgtcttcagacaagagccgatgcctgatcgagcagagcacgtgaacaattctatagtgcgacatccaatcgcgccgg
gcatctgtgccttaccagaccagacaaactctccagaagagcccggagcgtggcctcccagtctatcatgcctatacatgtccctg
ggcgccgagaacagcgtggcctacttaacaatagcatcgccatcccaaccaacttcaaatctctgtgaccacagagatcctgcccggt
ccatgaccaagacatctgtgactgcacaatgtatatctgtggcgattctaccgagtgagcaacctgctgctccagtacggcagcttttga
cccagctgaatagagccctgacaggcatcgccgtggagcaggataagaacacacaggaggtgtcgccaggtgaagcaaatctaca
gacccccctatcaaggactttggcggttc[caa]tttccagatcctgctgatccatccaagccttctaagcggagctttatcgaggacct
gctgttaacaaggtgacctggccgatcgccgcttcataagcagtatggcgattgcctggcgacatcgagccagggaacctgatctg
cgcccagaagtttaatggcctgacctgctgccaccctgctgacagatgagatgatcgacagtacacaagcgccctgctggccggcac
catcatccggatggacctcgcgaggagccgcccctccagatcccccttgccatgcagatggcctataggttcaacggcatcgcgctg
accagaatgtgctgtacgagaaccagaagctgatcgccaatcagtttaactccgccatcggaagatccaggacagcctgtcctctacag
ccagcgccctgggaagctccaggtgtggtgaatcagaaccccaggccctgaataacctggtgaagcagctgagcagcaacttcgg
cgccatctctagcgtgctgaatgacatcctgagccggctggacaaggtggaggcagaggtgcagatcgacggctgatcaccggccgg
ctccagagcctccagacctatgtgacacagcagctgatcagggccggcagatcagggccagcgccaatctggcagcaaccaagatgt

ccgagtgctgctgggcccagctctaagagagtggtgactttgtggcaagggtatcacctgatgtccttcctcagctgccccacacggcgtg
 gtgtttctgcacgtgacctacgtgcccggccaggagaagaacttcaccacagcccctgccatctgccacgatggcaaggcccactttccaa
 gggagggcgtgttcgtgtccaacggcaccactggtttgtgacacagcgcaatttctacgagccccagatcatcaccacagacaacacctt
 cgtgagcggcaactgtgacgtggtcatcggtcatcgtgaacaataccgtgtatgatccactccagcccagctggacagctttaaggagga
 gctggataagtatttcaagaatcacacctcccctgacgtggatctggcgacatcagcggcatcaatgcctccgtggtgaacatccagaag
 gagatcgaccgcctgaacgaggtggctaagaatctgaacgagagcctgatcgacctccaggagctgggcaagtatgagcagtacatcaa
 gtggccctgttacatctggctgggtctcatcgccggcctgatcgccatcgtgatggtgacctcatgctgtgctgtatgacatcctgctgttct
 tgcctgaagggtgctgtagctgtggctcctgctgtaagttgacgaggatgactctgaacctgtgctgaaggcgctgaagctgcattacac
 ctaa (SEQ ID NO: 9)

[00195] Protein sequence of S-(deg-RBD-801) for Wuhan strain (from N-terminus to C-terminus):

MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFS
 NVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTKSNIRGWIFGTTLDLSTQSLIV
 NNATNVVIKVCFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFMD
 LEGKQGNFKNLREFVFKNIDGYFKIYKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQ
 TLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLS
 ETKCTLKSTVEKGIYQTSNFRVQPAEIVRFPQITNLCPFGEVFAIRFASVYAWNRRK
 RISNCVADYSVLVNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVQRQIAPGQTG
 KIADYNYKLPDDFTGCVIAWNSNNLDSKVGNNYNYLYRLFRKSNLKPFERDISTEYIQA
 GSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVLSFELLHAPATVCGPKKSTNL
 VKNKCVNFNENGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFG
 GVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGNSNVFQTRAGCLIG
 AEHVNNSECDIPGAGICASYQTQTNPRRARSVASQSIIAYTMSLGAENSVAYSNNSI
 AIPTNFTISVTTEILPVSMKTSTVDCITMYICGDSTECNNLLQYGSFCTQLNRALTGIAVE
 QDKNTQEVFAQVKQIYKTPPIKDFGGFQFSQILPDPSKPSKRSFIEDLLFNKVTADAGFI
 KQYGDCLGDIAARDLCAQKFENGLTVLPPLTDEMIAQYTSALLAGTITSGWTFGAGAA
 LQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSTASALGKLQDVV
 NQNAQALNTLVKQLSSNFGAISSVLNDILSRLDPPEAEVQIDRLITGRLQSLQTYVTQQLI
 RAAEIRASANLAATKMSECVLGQSKRVDFCGGYHLSFPQSAPHGVVFLHVTYVPA
 QEKNTTAPAICHGDKAHFPREGVVFVSNGTHWFVTQRNFYEPQIITDNTFVSGNCDVV
 IGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVLGDISGINASVVNIQKEIDRLNEV
 AKNLNESLIDLQELGKYEYIKWPWYIWLGFIAGLIAVMVTIMLCCMTSCCSCLKGCC
 SCGSCCKFDEDDSEPVKGVKLHYT (SEQ ID NO: 10)

[00196] DNA or RNA sequence of S-(deg-RBD-801) for Wuhan S-2P strain (from 5'-end to 3'-end):

atgttcgtcttcttggtcctgctgcctctggtctcctcacagtgcgtcaatctgacaactcggactcagctgccacgtgcttataactaatagcttc
accagaggcgtgtactatcctgacaaggtgttagaagctccgtgctgcactctacacaggatctgtttctgccattcttagcaacgtgacct
ggttcacgcatccacgtgagcggcaccaatggcacaagcggctgacaatcccgtgctgcctttaacgatggcgtgtacttcgcctct
accgagaagagcaacatcatcagaggctggatcttggcaccacactggactccaagacacagtctctgctgatcgtgaacaatgccacc
aacgtggtcatcaaggtgtgcgagttccagtttgaatgatccctcctgggcgtgtactatcacaagaacaataagagctggatggagtcc
gagtttagagtgtattctagcgccaacaactgcacatttgagtacgtgagccagccttctgatggacctggaggcgcaagcagggaattt
caagaacctgagggagttcgtgtttaagaatatcgacggctacttcaaatctactctaagcacacccccatcaacctggtgcgcgacctgc
ctcagggcttcagcgccctggagcccctggtggatctgcctatcgccatcaacatcacccggttcagacactgctggccctgcacagaag
ctacctgacacccggcgactcctctagcggatggaccgcccggcgtgcccctactatgtgggctacctccagccccggaccttctgct
gaagtacaacgagaatggcaccatcacagacgcagtgattgcgcctggacccccctgagcgagacaaagtgtactgaagtctttac
cgtggagaaggcatctatcagacatccaatttcagggtgcagccagccgagcgatcgtgcgcttctgaaatcacaaacctgtgccca
tttggcgagggtgtcagggaacccgcttcgacgcgtgtacgcctggaataggaagcggatcagcaactgcgtggccgactatagcgtg
ctgtacaactccgctctttcagcacctttaagtctatggcgtgtccccacaaagctgaatgacctgtgctttaccaacgtctacgccgattc
ttctgctgatcagggcgacgaggtgcgccagatcgccccggccagacaggcaagatcgagactacaattataagctgccagacgattt
caccggctgcgtgatcgctggaacagcaacaatctggattccaaagtggcgccgcaactacaattatctgtaccggctgttagaaagagc
aatctgaagcccttcgagaggacatctctacagaaatctaccaggccggcagcaccccttgcaatggcgtggagggtttaactgttattt
cccactccagtcctacggcttcagccccacaaacggcgtgggctatcagccttaccgctggtggtgctgagctttgagctgctgcacgcc
ccagcaacagtgtgcggcccaagaagtccaccaatctggtgaagaacaagtgcgtgaactcaactcaacggcctgaccggcacagg
cgtgctgaccgagtcacaagaagttcctgccatttcagcagttcggcaggacatcgagataccacagacgccgtgcgcgaccac
agacctggagatcctggacatcacacctgctcttctcgcgcgctgagcgtgatcacacccggccaccaatacaagcaaccaggtggcc
gtgctgtatcaggacgtgaattgtaccgaggtgcccgtggctatccacgccgatcagctgaccccaacatggcgggtgtacagcaccggc
tccaacgtcttcagacaagagccgatcctgatcgagcagagcacgtgaacaattcctatgagtgcgacatcccaatcgccgcccgc
atctgtgctcttaccagaccagacaaactctcccagaagagcccggagcgtggcctcccagctctatcatcgctataccatgtccctggg
cgccgagaacagcgtggcctactctaacaatagcatcgcatcccaaccaacttcacaatctctgtgaccacagagatcctgccgtgtcc
atgaccaagacatctgtggactgcacaatgtatatctgtggcgattctaccgagtgagcaacctgctgctccagtacggcagcttttacc
cagctgaatagagccctgacaggcatcgccgtggagcaggataagaacacacaggaggtgttcgccaggtgaagcaaatctacaaga
cccccttatcaaggactttggcggttccaaatctccagatcctgctgatccatccaagccttctaagcggagctttatcaggacgtgct
gttcaacaaggtgacctggccgatccggcttcatcaagcagatggcgattgcctggcgacatcgagccaggacctgatctgcgc
ccagaagttaattggcctgacctgctgccacccctgctgacagatgagatgatcgacagtacacaagcgcctgctggccggcaccat
cacatccggatggaccttcggcgaggagccgcccctccagatcccccttgcctgagatggcctataggttaacggcatcgcgctgac
ccagaatgtgctgtacgagaaccagaagctgatcgcaatcagtttaactccgccatcggaagatccaggacagcctgtcctctacagcc
agcgccctgggcaagctccaggatgtgtgaatcagaacgcccaggccctgaataccctggtgaagcagctgagcagcaacttcggcg

ccatctctagcgtgctgaatgacatcctgagccggctggacccgcccggaggcagaggtgcagatcgaccggctgacaccggccggct
ccagagcctccagacctatgtgacacagcagctgatcagggccgcccagatcagggccagcgccaatctggcagcaaccaagatgtcc
gagtgctgctgggcccagctaaagagagtggaactttgtggcaagggtatcacctgatgccttcctcagctgccccacacggcgtggt
gtttctgcacgtgacctacgtgccccccaggagaagaacttcaccacagcccctgcatctgccacgatggcaaggcccactttccaag
ggagggcgtgttctgtccaacggcaccactggttgtgacacagcgcaatttctacgagccccagatcatcaccacagacaacaccttc
gtgagcggcaactgtgacgtggtcatcgcatcgtgaacaataccgtgtatgatccactccagcccagctggacagctttaaggaggag
ctggataagtatttcaagaatcacacctcccctgacgtggatctggggacatcagcgcatcaatgcctccgtggtgaacatccagaagg
agatcgaccgctgaacgaggtggctaagaatcgaacgagacgtgatcgacctccaggagctgggcaagtatgagcagtacatcaag
tgccctggtacatctggctgggttcatcgccggcctgatcgccatcgtgatggtgacctcatgctgtgctgatgacatctgctgttctt
gcctgaagggtgctgtagctgtggctcctgctgtaagttgacgaggatgactctgaacctgtgctgaagggcgtgaagctgcattacacc
taa (SEQ ID NO: 35)

[00197] **Protein sequence of S-(deg-RBD-801) for Wuhan S-2P strain (from N-terminus to C-terminus):**

MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFS
NVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTKSNIRGWIFGTTLDSTQSLIV
NNATNVVIKVFCEQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMD
LEGKQGNFKNLREFVFNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQ
TLLALHRSYLTPGDSSSGWTAGAAAYVGYLQPRTFLLKYNENGTITDAVDCALDPLS
ETKCTLKSFTVEKGIYQTSNFRVQPAEAIVRFPQITNLCPFGEVFQATRFASVYAWNKRK
ISNCVADYSVLVNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRIAPGQTG
KIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNLYRLFRKSNLKPFERDISTEYIQA
GSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVLSFELLHAPATVCGPKKSTNL
VKNKCVNFNENGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSF
GVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIG
AEHVNNSYECDIPIGAGICASYQTQTNSPRRARSVASQSIIAYTMSLGAENSVAYSNSI
AIPTNFTISVTTEILPVSMKTSTVDCITMYICGDSTECNLLLQYGSFCTQLNRALTGIAVE
QDKNTQEVFAQVKQIYKTPPIKDFGGFQFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFI
KQYGDCLGDIAARDLICAQKFENGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAA
LQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSTASALGKLQDVV
NQNAQALNTLVKQLSSNFGAISSVLNDILSRDPEAEVQIDRLITGRLQSLQTYVTQQLI
RAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPA
QEKNTTAPAICHGDKAHFPREGVVFVSNGTHWFTVQRNFYEPQIITDNTFVSGNCDVV
IGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVLDISGINASVVNIQKEIDRLNEV

AKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGLIAIVMTIMLCCMTSCCSCLKGCC
SCGSCCKFDEDDSEPVLKGVKLHYT (SEQ ID NO: 36)

[00198] DNA or RNA sequence of S-(deg-RBD-1194) for Wuhan strain (from 5'-end to 3'-end):

atgttcgtcttctggctcctgctgcctctggtctcctcacagtgcgtcaatctgacaactcggactcagctgccacctgcttataactaatagcttc
accagaggcgtgtactatcctgacaaggtgttagaagctccgtgctgcactctacacaggatctgttctgccattcttagcaacgtgacct
ggttcacgccatccacgtgagcggcaccatggcacaagcggctcgacaatcccgtgctgccttttaacgatggcgtgtacttcgcctct
accgagaagagcaacatcatcagaggctggatcttggcaccacactggactccaagacacagtctctgctgatcgtgaacaatgccacc
aacgtggtcatcaaggtgtgcgagttccagtttgaatgatccctcctggcgctgtactatcacaagaacaataagcgtggatggagtc
gagtttagagtgtattctagcgccaacaactgcacatttgagtacgtgagccagccttctgatggacctggagggaagcagggaattt
caagaacctgagggaagtctgttgaagaatcgcaggctacttcaaatctacttaagcacacccccatcaacctggtgcgcgacctgc
ctcagggcttcagcgccctggagcccctgggtggtatcgcctatcgcatcaacatcacccgggttcagacactgctggccctgcacagaag
ctacctgacacccggcgactcctctagcggatggaccgcccggcgctgccctactatgtgggctacctccagccccggaccttctgct
gaagtacaacgagaatggcaccatcacagacgcagtggtgctgcctggacccctgagcgcagacaaagtgtactgaagtccttac
cgtggagaaggcatctatcagacatccaatttcagggtgcagccagccgagcgatcgtgcgcttctcgaatcacaaacctgtgcc
atttggcgaggtgttcgagcgaacccgcttcgcagcgtgtacgcctggaataggaagcggatcagcaactgcgtggccgactatagcgt
gctgtacaactccgctcttctcagcacctttaaagtctatggcgtgtccccacaaagctgaatgacctgtgctttaccaacgtctacgccgat
tcttctgctgatcagggcgacgaggtgcgcagatcgccccggccagacaggcaagatcgagactacaattataagctgccagacga
tttaccggctgcgtgatcgctggaacgaacaatctggattccaaagtggcgccgaactacaattatctgtaccggctgtttagaaga
gcaatctgaagcccttcgagaggacatctctacagaaatctaccaggccggcagcacccttgcaatggcgtggagggtttaactgtta
tttccactccagtcctacggcttcagcccaaaacggcggtggctatcagccttaccgctgggtgctgagctttgagctgctgcacg
ccccagcaacagtgctgggccccagaagtccaccaatctgggtgaagaacaagtgcgtgaactcaacttaacggcctgaccggcaca
ggcgtgctgaccgagtcacaagaagtctcctgccaattcagcagttcggcaggacatcgagataccacagacggcgtgcgcgaccc
acagaccctggagatcctggacatcacacctgctcttccggcgcgctgagcgtgatcacacccggcaccataaagcaaccagggtg
ccgtgctgtatcaggacgtgaattgtaccgaggtgcccgtggctatccacgccgatcagctgacccaacatggcgggtgtacgacccg
gctccaacgtcttccagacaagagccgatgcctgatcgagcagagcacgtgaacaattcctatgagtcgacatcccaatcggcgcgc
gcatctgtgcctcttaccagaccagacaaactctccagaagagcccgagcgtggcctccagctctatcatcgctatacatgtccctg
ggcgccgagaacagcgtggcctacttaacaatagcatcgccatcccaacttcacaatctctgtgaccacagagatcctgcccgtgt
ccatgaccaagacatctgtgactgcacaatgtatatctgtggcgattctaccgagtgagcaacctgctgctccagtacggcagctttgt
cccagctgaatagagccctgacaggcatcgccgtggagcaggataagaacacagagggtgtcgccaggtgaagcaaatctacaa
gacccccctatcaaggactttggcggttcaattttccagatcctgcctgatccatcaagccttctaagcggagctttatcaggacctg
ctgttcaacaagtgacctggccgatcgccgttcatcaagcagtatggcgattgctggcgacatcgagccaggacctgatctgc
gcccagaagttaatggcctgacctgctgccacccctgctgacagatgagatgatgcacagtacacaagcgccctgctggccggcacc
atcacatccggatggaccttcggcgagcagccgcccctccagatccccttgcctgagatggcctataggttcaacggcatcgccgtga

cccagaatgtgctgtacgagaaccagaagctgatcgccaatcagtttaactccgccatcggaagatccaggacagcctgtcctctacagc
cagcgccctgggcaagctccaggatgtggtgaatcagaacgcccaggccctgaataccctgggaagcagctgagcagcaacttcggc
ggcatctctagcgtgctgaatgacatcctgagccggctggacaaggtggaggcagaggtgcagatcgaccggctgatcaccggccggct
ccagagcctccagacctatgtgacacagcagctgatcaggggccgagatcagggccagcgccaatctggcagcaaccaagatgtcc
gagtgcgtgctgggcccagctaaagagagtggaactttgtggcaagggtgatcacctgatgtccttcctcagctgtccccacacggcgtggt
gtttctgcacgtgacctacgtgcccggccaggagaagaacttcaccacagcccctgccatctgccacgatggcaaggcccactttccaag
ggagggcggtgtctgtccaacggcaccactggtttgtgacacagcgcaattttacgagccccagatcatcaccacagacaacaccttc
gtgagcggcaactgtgacgtggtcatcggcacgtgaacaataccgtgtatgatccactccagcccagctggacagctttaaggaggag
ctggataagtatttcaagaatcacacctcccctgacgtggatctggggacatcagcgcatcaatgctccgtggtgaacatccagaagg
agatcgaccgctgaacgaggtggctaagaatctgcaggagagcctgatcgacctccaggagctgggcaagtatgagcagtacatcaa
gtggccctgtacatctggctgggttcacgcccggcctgatcgccatcgtgatggtgaccatcatgctgtgctgtatgacatcctgtgttt
tgcctgaagggtgctgtagctgtggctcctgctgtaagttgacgaggatgactctgaacctgtgctgaagggtggaagctgcattacac
ctaa (SEQ ID NO: 11)

[00199] Protein sequence of S-(deg-RBD-1194) for Wuhan strain (from N-terminus to C-terminus):

MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFS
NVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTKSNIRGWIFGTTLDLSDKTSQSLIV
NNATNVVIKVFCEQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMD
LEGKQGNFKNLREFVFKNIDGYFKIYKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQ
TLLALHRSYLTTPGDSSSGWTAGAAAYVGYLQPRTFLLKYNENGTITDAVDCALDPLS
ETKCTLKSFTVEKGIYQTSNFRVQPAEAIIVRFPQITNLCPFGEVFQATRFASVYAWNRRK
RISNCVADYSVLVNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVQRQIAPGQTG
KIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNLYRLFRKSNLKPFERDISTEYIQA
GSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVLSFELLHAPATVCGPKKSTNL
VKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFQ
GVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIG
AEHVNNSYECDIPIGAGICASYQTQTNPRRARSVASQSIIAYTMSLGAENSVAYSNNSI
AIPTNFTISVTTEILPVSMKTSTVDCITMYICGDSTECNLLLQYGSFCTQLNRALTGIAVE
QDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTADAGFI
KQYGDCLGDIAARDLCAQKFNGLTVLPLLDEMIQYTSALLAGTITSGWTFGAGAA
LQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSTASALGKLQDVV
NQNAQALNTLVKQLSSNFGAISSVLNDILSRLDPPEAEVQIDRLITGRLQSLQTYVTQQLI
RAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSPQSAPHGVVFLHVTYVPA

QEKNTTAPAICHGKAHFPREGVFVSNGTHWFTVQRNFYEPQIITDNTFVSGNCDVV
 IGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVLDISGINASVVNIQKEIDRLNEV
 AKNLQESLIDLQELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCCLKGCC
 SCGSCCKFDEDDSEPVLKGVKLHYT (SEQ ID NO: 12)

[00200] DNA or RNA sequence of S-(deg-RBD-1194) for Wuhan S-2P strain (from 5'-end to 3'-end):

atgttcgtcttctggtcctgctgcctctggtctcctcacagtgcgtcaatctgacaactcggactcagctgccacctgcttataataatagcttc
 accagaggcgtgtactatcctgacaagggtttagaagctccgtgctgactctacacaggatctgtttctgccattcttagcaacgtgacct
 ggttcacgccatccacgtgagcggcaccaatggcacaagcgggtcgacaatcccgtgctgcttttaacgatggcgtgtacttcgcctct
 accgagaagagcaacatcatcagaggctggatcttggcaccacactggactccaagacacagtctctgctgatcgtgaacaatgccacc
 aacgtggtcatcaagggtgctgcgagttccagtttgaatgatccctcctggcgtgtactatcacaagaacaataagagctggatggagtc
 gaggtttagtgattctagcgccaacaactgcacatttgagtacgtgagccagccttctgatggacctggagggaagcagggaattt
 caagaacctgagggagttcgtgttgaagaatatcgacggctactcaaatctactctaagcacacccccatcaacctggtgcgcgacctgc
 ctgagggcttcagcgccctggagcccctggtggtatgcctatcgcatcaacatcacccgggttcagacactgctggccctgcacagaag
 ctacctgacacccggcgactcctctagcggatggaccgcccggcgtgccgctactatgtgggctacctccagccccggaccttctgct
 gaagtacaacgagaatggcaccatcacagacgcagtggtggtgcgcctggacccccctgagcgagacaaagtgtactgaagtcctttac
 cgtggagaaggcatctatcagacatccaatttcagggtgcagccagccgagcgatcgtgcgcttctgaaatcacaaacctgtgcca
 tttggcgagggtgttcgaggcaacccgcttcgcccagcgtgtacgcctggaataggaagcggatcagcaactgcgtggccgactatagcgtg
 ctgtacaactccgctctttcagcaccttaagtctatggcgtgtccccacaaagctgaatgacctgtgctttaccaacgtctacgccgattc
 tttcgtgatcagggcgacgaggtgcgcagatcgccccggccagacaggcaagatcgagactacaattataagctgccagacgattt
 caccggctgcgtgatcgctggaacagcaacaatctggattccaaagtggcgcgcaactacaattatctgtaccggctgtttagaagagc
 aatctgaagcccttcgagaggacatctctacagaaatctaccaggccggcagcacccttgcaatggcgtggagggttttaactgttattt
 cccactccagctctacggctccagcccacaaacggcgtgggctatcagccttaccgctggtggtgctgagctttgagctgctgcacgcc
 ccagcaacagtgctgcggcccaagaagtccaccaatctggtgaagaacaagtcgtgaactcaactcaacggcctgaccggcacagg
 cgtgctgaccgagtcacaagaagtctcctgcccatttcagcagttcggcagggacatcgagataccacagacggcgtgcgcgaccac
 agacctggagatcctggacatcacacctgctcttctggcggcgtgagcgtgatcacacccggcaccaatacaagcaaccaggtggcc
 gtgctgtatcaggacgtgaattgtaccgaggtgcccgtggctatccacgccgatcagctgaccccaacatggcgggtgtacagcaccggc
 tccaacgtcttcagacaagagcgggatgcctgatcgagcagagcacgtgaacaattcctatgagtgcgacatccaatcgccgcccgc
 atctgtgcctcttaccagaccagacaaactctcccagaagagcccggagcgtggcctcccagctctatcatcgccctatacatgtccctggg
 cgccgagaacagcgtggcctactctaacaatagcatcgccatcccaaccaacttcacaatctctgtgaccacagagatcctgcccgtgtcc
 atgaccaagacatctgtggactgcacaatgtatatctgtggcgattctaccgagtgagcaacctgctgctccagtacggcagctttgtacc
 cagctgaatagagccctgacaggcatcgccgtggagcaggataagaacacacaggaggtgttcgccaggtgaagcaaatctacaaga
 cccccctatcaaggactttggcggcttcaattttccagatcctgcctgatccatccaagccttctaagcggagctttatcaggagacctgct

gttcaacaaggtgacctggccgatgccggcttcatcaagcagtagtgccgattgcctggcgacatcgagccaggacatgatctgcgc
ccagaagttaatggcctgacctgctgccaccctgctgacagatgagatgatcgacagtagacacaagcgccctgctggccggcaccat
cacatccggatggaccttggcgaggagccgacctccagatcccccttggccatgagatggcctataggttcaacggcatcgccgtgac
ccagaatgtgctgtacgagaaccagaagctgatcgccaatcagtttaactccgccatcggaagatccaggacagcctgtcctctacagcc
agcgccctgggcaagctccaggatgtgtgaatcagaacgcccaggccctgaataacctggtgaagcagctgagcagcaacttggcg
ccatctctagcgtgctgaatgacatcctgagccggctggacccgcccggagggcagaggtgcagatcgaccggctgatcaccggccggct
ccagagcctccagacatgtgacacagcagctgatcaggccgcccagatcaggccagcgccaatctggcagcaaccaagatgtcc
gagtgctgtctgggcccagctaaagagtggaactttgtggcaagggtatcacctgatgtccttccctcagctgtccccacacggcgtggt
gtttctgcacgtgacctacgtgccgcccaggagaagaacttaccacagccccctgccatctgccagatggcaaggcccactttcaag
ggagggcgtgttctgtccaacggcaccactggtttgtgacacagcgcaatttctacgagccccagatcatcaccagacaacaccttc
gtgagcggaactgtgacgtggtcatcggtatgaacaataacctgtatgatccactccagcccagctggacagctttaaggaggag
ctggataagtattcaagaatcacacctccccctgacgtggtatctggcgacatcagcgcatcaatgctccgtggtgaacatccagaagg
agatcgaccgctgaacgaggtggctaagaatctgaggagagcctgatcgacctccaggagctgggcaagtatgagcagtagcatcaag
tggccctggtacatctggtggcttcatcgccggcctgatcgccatcgtgatggtgacctcatgctgtgctgtatgacatcctgctgttctt
gcctgaagggtgctgtagctgtggctcctgctgtaagttgacgaggatgactctgaacctgtgctgaaggcgctgaagctgcattacacc
taa (SEQ ID NO: 37)

[00201] Protein sequence of S-(deg-RBD-1194) for Wuhan S-2P strain (from N-terminus to C-terminus):

MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFS
NVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIV
NNATNVVIVKCEFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMD
LEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQ
TLLALHRSYLTGPDSSSGWTAGAAAYVGYLQPRTFLLKYNENGTITDAVDCALDPLS
ETKCTLKSFTVEKGIYQTSNFRVQPAEAIVRFPQITNLCPFGEVFQATRFASVYAWNRRK
ISNCVADYSVLVNSASFSTFKCYGVSPKLNLDLCTNVYADSFVIRGDEVRIAPGQTG
KIADYNYKLPPDDFTGCVIAWNSNNLDSKVGGNYNLYRLFRKSNLKPFERDISTEYIQA
GSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVLSFELLHAPATVCGPKKSTNL
VKNKCVNFNFNGLTGTGVLTESNKKFLPFQGFGRDIADTTDAVRDPQTLEILDITPCSFG
GVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIG
AEHVNNSECDIPIGAGICASYQTQTNPRRARSVASQSIIAYTMSLGAENSVAYSNNNSI
AIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVE
QDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFI
KQYGDCLGDIAARDLCAQKFNGLTVLPLLTDemiaQYTSALLAGTITSGWTFGAGAA
LQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSTASALGKLQDVV

NQNAQALNTLVKQLSSNFGAISSVLNDILSRDPPAEVQIDRLITGRLQSLQTYVTQQLI
 RAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPA
 QEKNFTTAPAICHGDKAHFPREGV FVSNGTHWFVTQRNFYEPQIITDNTFVSGNCDVV
 IGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVLGDISGINASVVNIQKEIDRLNEV
 AKNLQESLIDLQELGKYEQYIKWPWYIWLGFIAGLIAIVMTIMLCCMTSCCSCLKGCC
 SCGSCCKFDEDDSEPVLKGVKLHYT (SEQ ID NO: 38)

[00202] DNA or RNA sequence of S-(deg-RBD-122-165-234) for Wuhan strain (from 5'-end to 3'-end):

atgttcgtcttcttggtcctgctgcctctggtctcctcacagtgcgtcaatctgacaactcggactcagctgccacgtgcttataactaatagcttc
 accagaggcgtgtactatcctgacaaggtgttagaagctccgtgctgactctacacaggatctgtttctgccattcttagcaacgtgacct
 ggttcacgccatccacgtgagcggcaccaatggcacaagcgggtcgacaatcccgtgctgcttttaacgatggcgtgtacttcgctct
 accgagaagagcaacatcatcagaggctggatcttggcaccacactggactccaagacacagtctctgctgatcgtgaac[ca]gccacc
 aacgtggtcatcaaggtgtgcgagttccagttttgtaatgatccctcctggcgtgtactatcacaagaacaataagagctggatggagtc
 gagtttagagtgtattctagcgccaac[ca]gtgcacattgagtacgtgagccagcctttcctgatggacctggagggaagcagggaattt
 caagaacctgagggagtgctgtttaagaatatcgacggctacttcaaatctacttaagcacacccccatcaacctggtgcgcgacctgc
 ctgagggtctcagcgccctggagcccctggtggatctgcctatcggcatc[ca]gtacacccggttcagacactgctggccctgcacagaa
 gtactctgacacccggcgtactccttagcggatggaccgcccggcgtgccgcctactatgtgggctacctccagccccggaccttctgc
 tgaagtacaacgagaatggcaccatcacagacgcagtggtggtgcgcctggaccctgagcgagacaaagtgtactgaagtccttta
 ccgtggagaagggtcatctatcagacatccaattcagggtgcagcca[cc]gag[cc]gagcgtgcgcttct[ga]atcacaacctgtgcc
 catttggcgaggtgttc[ga]gcaacccgcttcgccagcgtgtacgcctggaataggaagcggatcagcaactgcgtggccgactatagc
 gtgctgtacaactccgctcttccagcaccttaagtgtatggcgtgtccccacaaagctgaatgacctgtgctttaccaactctacgccg
 attcttctgtgatcagggcgacgaggtgcgacagatcgccccggccagacaggcaagatcgagactacaattataagctgccagac
 gatttaccggctgcgtgatcgctggaacagcaacaatctggattccaaagtggcggaactacaattatctgtaccggctgttagaaa
 gagcaatctgaagccctcgagaggacatctctacagaatctaccaggccggcagcacccttgcaatggcgtggagggtttaaactgt
 tatttccactccagtcctacggcttccagcccacaaacggcgtgggctatcagccttaccgctggtggtgctgagctttgagctgctgcac
 gccccagcaacagtgtgcggccccaagaagtccaccaatctggtgaagaacaagtgcgtgaactcaactcaacggcctgaccggcac
 aggcgtgctgaccgagtcacaagaagttcctgccatttcagcagttcggcagggacatcgagataccagacgcccgtgcgcgacc
 cacagaccctggagatcctggacatcacaccctgctcttccggcggcgtgagcgtgatcacacccggcaccataagaacagaggtg
 gccgtgctgtatcaggacgtgaattgtaccgaggtgcccgtggctatccacgccgatcagctgacccaacatggcgggtgtacagcacc
 ggctccaacgtcttccagacaagagccggatgcctgatcggagcagagcacgtgaacaattcctatgagtgcgacatcccaatcggcgcc
 ggcatctgtgcctcttaccagaccagacaaactctccagaagagccggagcgtggcctcccagctatcatcgccctatcatatgtccct
 gggcgcggagaacagcgtggcctacttaacaatagcatcgccatcccaaccaacttcacaatctgtgaccacagagatcctgccgtg
 tccatgaccaagacatctgtggactgcacaatgtatatctgtggcgattctaccgagtgacgaacctgctgctccagtacggcagctttgta
 ccagctgaatagagccctgacaggcatcgccgtggagcaggataagaacacagggaggtgtcgccagggtgaagcaaatctacaa

gacccccctatcaaggactttggcggctcaattttccagatcctgcctgatccatccaagccttctaagcggagctttatcgaggacctg
ctgttcaacaaggtagacctggccgatgccggctcatcaagcagtagggcattgcctggcgacatcgagccaggacatgatctgc
gcccagaagttaatggcctgacctgctgccacccctgctgacagatgagatgatcgacagtacacaagcgccctgctggccggcacc
atcacatccggatggaccttcggcgaggagccgcccctccagatccccttggcatgcagatggcctataggttaacggcatcgccgtga
cccagaatgtgctgtacgagaaccagaagctgatcgccaatcagtttaactccgccatcggaagatccaggacagcctgtcctctacagc
cagcgccctgggcaagctccaggatgtggtgaatcagaacgccaggccctgaataccctggtgaagcagctgagcagcaacttcggc
ggcatctctagcgtgctgaatgacatcctgagccggctggacaagggtggaggcagaggtgcagatcgaccggctgatccggccggct
ccagagcctccagacctatgtgacacagcagctgatcaggccggccgagatcaggccagcgccaatctggcagcaaccaagatgtcc
gagtgcgtgctggccagctaaagagtgacttttgggaagggtatcacctgatgccttcctcagctgtccccacagcgctggt
gtttctgcacgtgacctacgtccccggcaggagaagaacttcaccacagcccctgccatctgccacgatggcaaggcccactttccaag
ggagggcggttctgtccaacggcaccactggtttgtgacacagcgcaattttacgagccccagatcatcaccacagacaacaccttc
gtgagcggcaactgtgacgtggtcatcgccatcgtgaacaataccgtgtatgatccactccagcccagctggacagctttaaggaggag
ctggataagtattcaagaatcacacctcccctgacgtggatctggcgacatcagcgccatcaatgcctccgtggtgaacatccagaagg
agatcgaccgcctgaacgaggtggctaagaatcgaacgagcctgatcgacctccaggagctgggcaagtatgagcagtagcatcaag
tggccctgtgacatctggctggcctcatcgccggcctgatcgccatcgtgatggtgacctcatgctgtgctgtatgacatcctgctgttctt
gcctgaagggtgctgtagctgtggtcctgctgtaagttgacgaggatgactctgaacctgtgctgaaggcgctgaagctgcattacacc
taa (SEQ ID NO: 13)

[00203] Protein sequence of S-(deg-RBD-122-165-234) for Wuhan strain (from N-terminus to C-terminus):

MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFS
NVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTKSNIRGWIFGTTLDLSTQSLIV
N^QATNVVIKVFCEQFCNDPFLGVYYHKNNKSWMESEFRVYSSAN^QCTFEYVSQPFLM
DLEGKQGNFKNLREFVFNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLP^QITRF
QTLALHRSYLTSGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPL
SETKCTLKSFTVEKGIYQTSNFRVQPA^E^EIVRFP^QITNLCPFGEVF^QATRFASVYAWNR
KRISNCVADYSVLNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRIAPGQT
GKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIQ
AGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVLSFELLHAPATVCGPKKSTN
LVKNKCVNFNFNGLTGTGVLTESNKKFLPFQFGRDIADTTDAVRDPQTLEILDITPCSF
GGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLI
GAEHVNNSYECDIPIGAGICASYQTQTNSPRRARSVASQSIAYTMSLGAENSVAYSNNNS
IAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECNLLLQYGSFCTQLNRALTGIAVE
QDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFI
KQYGDCLGDIAARDLICAQKFNGLTVLPLLTDemiaQYTSALLAGTITSGWTFGAGAA

LQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSTASALGKLQDVV
NQNAQALNTLVKQLSSNFGAISSVLNDILSRDPPEAEVQIDRLITGRLQSLQTYVTQQLI
RAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSPQSPHGVVFLHVTYVPA
QEKNTTAPAICHGDKAHFPREGVVFVSNGTHWFTQRFYEPQIITDNTFVSGNCDVV
IGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVLDGDISGINASVUNIQKEIDRLNEV
AKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCLKGCC
SCGSCCKFDEDDSEPVLKGVKLHYT (SEQ ID NO: 14)

[00204] DNA or RNA sequence of S-(deg-RBD-122-165-234) for Wuhan S-2P strain (from 5'-end to 3'-end):

atgttcgtcttcctggctcctgctgcctctggtctcctcacagtgcgtcaatctgacaactcggactcagctgccacctgcttatactaataagcttc
accagaggcgtgtactatcctgacaaggtgttagaagctccgtgctgcactctacacaggatctgtttctgccattcttagcaacgtgacct
ggttcacgcatccacgtgagcggcaccaatggcacaagcggctgacaatcccgtgctgcctttaacgatggcgtgtacttcgcctct
accgagaagagcaacatcatcagaggctggatcttggcaccacactggactccaagacacagtctctgctgatcgtgaaccaagccacc
aacgtggtcatcaaggtgtgcgagttccagtttgaatgatccctcctggcgctgtactatcacaagaacaataagagctggatggagtcc
gagtttagagtgtattctagcgccaaccagtgcacattgagtacgtgagccagccttctgatggacctggaggcgcaagcagggaattt
caagaacctgagggagttcgtgtttaagaatatcgacggctacttcaaatctactctaagcacacccccatcaacctggtgcgcgacctgc
ctcagggcttcagcgccctggagcccctggtggatctgcctatcgccatccagatcacccggttcagacactgctggccctgcacagaag
ctacctgacacccggcgactcctctagcggatggaccgcccggcgctgccgctactatgtgggctacctccagccccggaccttctgct
gaagtacaacgagaatggcaccatcacagacgcagtgattgcgccctggacccccctgagcgagacaaagtgtactgaagtcctttac
cgtggagaaggcatctatcagacatccaatttcagggtgcagccagccgagcgatcgtgcgcttctgaaatcacaaacctgtgcca
tttggcgagggtgttcgaggcaacccgcttcgccagcgtgtacgcctggaataggaagcggatcagcaactgcgtggccgactatagcgtg
ctgtacaactccgctcttccagcacctttaagtctatggcgtgtccccacaaagctgaatgacctgtgctttaccaacctctacgccgattc
ttcgtgatcagggcgacgaggtgcgccagatcgccccggccagacaggcaagatcgagactacaattataagctgccagacgattt
caccggctgcgtgatcgctggaacagcaacaatctggattccaaagtggcgaggcaactacaattatctgtaccggctgttagaaagagc
aatctgaagcccttcgagaggacatctctacagaaatctaccaggccggcagcacccttgcaatggcgtggagggtttaactgttattt
cccactccagtctacggttccagcccacaaacggcggtggctatcagccttaccgctggtggtgctgagctttgagctgctgcacgcc
ccagcaacagtgtgcggcccaagaagtccaccaatctggtgaagaacaagtgcgtgaactcaactcaacggcctgaccggcacagg
cgtgctgaccgagtcacaagaagttcctgccatttcagcagttcggcaggacatcgagataccacagacgccgtgcgcgaccac
agacctggagatcctggacatcacacctgctcttctggcggtgagcgtgatcacacccggcacaatacaagcaaccaggtggcc
gtgctgtatcaggacgtgaattgtaccgaggtgccgtggctatccacgccgatcagctgacccaacatggcgggtgtacagaccggc
tccaacgtcttcagacaagagccggatgcctgatcgagcagacgtgaacaattctctatgagtcgacatcccaatggcgccggc
atctgtgcctcttaccagaccagacaaactctcccagaagagcccgagcgtggcctcccagctctatcatcgctataccatgtccctggg
cgccgagaacagcgtggcctactctaacaatagcatcgccatcccaaccaacttcacaatctctgtgaccacagagatcctgccgtgtcc
atgaccaagacatctgtggactgcacaatgtatatctgtggcgattctaccgagtcgagcaacctgctgctccagtacggcagctttgtacc

cagctgaatagagccctgacaggcatgccgtggagcaggataagaacacacaggaggtgttcgccaggtgaagcaaatctacaaga
 cccccctatcaaggactttggcggctcaattttccagatcctgcctgatccatccaagccttctaagcggagctttatcgaggacctgct
 gttcaacaaggtgacctggccgatccggcttcatcaagcagatggcgattgcctggcgacatcgagccaggacctgatctgcc
 ccagaagttaatggcctgacctgctgccaccctgctgacagatgagatgatcgacagtacacaagcgcctgctggccggcaccat
 cacatccgatggaccttggcgagcggccctccagatcccccttggcatgagatggcctataggttaacggcatcggcgctgac
 ccagaatgtgctgtacgagaaccagaagctgatcgcaatcagtttaactccgccatcggaagatccaggacagcctgtcctctacagcc
 agcgccctgggcaagctccaggatgtgtgaatcagaacgccaggccctgaataacctggtgaagcagctgagcagcaactcggcg
 ccatctctagcgtgctgaatgacatcctgagccggctggacccggcgagggcagaggtgcagatcgaccggctgatcaccggccggct
 ccagagcctccagacctatgtgacacagcagctgatcaggcgccggagatcaggccagcgcaatctggcagcaaccaagatgtcc
 gagtgctgtctgggcccagctaaagagatggacttttgggaagggtatcacctgatgtccttccctcagctgtccccacacggcggtgt
 gtttctgcacgtgacctacgtgcccgcccaggagaagaactccacacagcccctgccatctgccagatggcaaggcccactttcaag
 ggagggcggtgtctgtccaacggcaccactggtttgtgacacagcgcaatttctacgagccccagatcatcaccagacaacaccttc
 gtgagcggaactgtgacgtggtcatcgccatcgtaacaatacgtgtatgatccactccagcccagctggacagctttaaggaggag
 ctggataagtatttaagaatcacacctccccctgacgtggatctggcgacatcagcgcatcaatgcctccgtggtgaacatccagaagg
 agatcgaccgcctgaacgaggtggctaagaatctgaacgagagcctgatcgacctccaggagctgggcaagtatgagcagtacatcaag
 tggccctggtacatctggctgggcttcatcgccggcctgatcgccatcgatggtgacctcatgctgtgtatgacatcctgctgttctt
 gcctgaagggtgctgtagctgtggctcctgctgtaagttgacgaggatgactctgaacctgtgctgaaggcgctgaagctgcattacac
 taa (SEQ ID NO: 39)

[00205] Protein sequence of S-(deg-RBD-122-165-234) for Wuhan S-2P strain (from N-terminus to C-terminus):

MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFS
 NVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSTQSLIV
 NQATNVVIKVCEFCNDPFLGVYYHKNNKSWMESEFRVYSSANQCTFEYVSQPFLMD
 LEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGIQITRFQ
 TLLALHRSYLTGPDSSSGWTAGAAAYVGYLQPRTFLLKYNENGTITDAVDCALDPLS
 ETKCTLKSFTVEKGIYQTSNFRVQPAEAIVRFPQITNLCPFGEVFQATRFASVYAWNKRK
 ISNCVADYSVLVNSASFSTFKCYGVSPKLNLCFTNVYADSFVIRGDEVRIAPGQTG
 KIADYNYKLPPDDFTGCVIAWNSNNLDSKVGGNYNLYRLFRKSNLKPFERDISTEYQA
 GSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVLSFELLHAPATVCGPKKSTNL
 VKNKC VN FN ENGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFG
 GVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIG
 AEHVNNSECDIPGAGICASYQTQTNPRRARSVASQSIIAYTMSLGAENSVAYSNNSI
 AIPTNFTISVTTEILPVSMTKTSVDCITMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVE
 QDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFI

KQYGDCLGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAA
 LQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSTASALGKLQDVV
 NQNAQALNTLVKQLSSNFGAISSVLNDILSRDPPPEAEVQIDRLITGRLQSLQTYVTQQLI
 RAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPA
 QEKNFTTAPAICHGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITDNTFVSGNCDVV
 IGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVLGDISGINASVVNIQKEIDRLNEV
 AKNL NESLIDLQELGKYEQYIKWPWYIWLGFIAGLIAIVMTIMLCCMTSCCSCCLKGCC
 SCGSCCKFDEDDSEPV LKGVKLHYT (SEQ ID NO: 40)

Table 1: The IG_{50} for variant pseudovirus neutralization assay.

IC ₅₀ serum dilution	WT	B.1.1.7	B.1.351	P1	B.1.617.2
WT	7658.4	1250.1	710.6	776.5	835.7
S-(deg-RBD)	6496.2	1926.5	1108.3	1523.6	1868.3
S-(deg-S2)	4825.6	4035.9	4568.3	4235.3	4786.3
S-(S2-1194)	5767.8	4628.6	5287.5	4989.6	5103.4

EXAMPLE 5: Glycosylation affected cellular and cytokine response

[00206] To characterize the T cell response, splenocytes from immunized mice were isolated and incubated with the peptide pool of S protein to measure the granzyme B (GrzB)-secreting T cells by elispot analysis. It was shown that S-(deg-S2) and S-(S2-1194) induced more GrzB-secreting cells than WT and S-(deg-RBD) did after incubation with full-length WT S (Fig. 7A), RBD (Fig. 7B) and S2 peptides (Fig. 7C), suggesting that glycosylation on S2 regulated the T cell response. To further study the effect on CD4⁺ and CD8⁺ T cells, the isolated T cells were incubated with bone-marrow-derived dendritic cells (DCs) and WT S peptide pool to measure the IFN γ -secreting T cells by flow cytometry. There was no significant difference in the number of the IFN γ -secreting CD4⁺ T cells among all vaccinated mice (Fig. 7D), but the mRNA vaccine of S-(deg-S2) or S-(S2-1194) induced more IFN γ -secreting CD8⁺ T cells, suggesting that glycosylation in S2 regulated CD8⁺ T cell response (Fig. 7E).

[00207] To analyze the cytokine expression, the medium from splenocytes incubated with full-length WT S peptide pool was measured by ELISA. It was shown that the splenocytes from S-(deg-S2) and S-(S2-1194) immunized mice secreted higher levels of T-helper-1 (TH1) cytokines (IFN γ , IL-2, and IL-12) (Fig. 8 A, B and E), whereas the splenocytes from WT and S-(deg-RBD) immunized mice secreted higher levels of T-helper-2 (TH2) cytokines (IL-4, IL-6 and IL-13) (Fig. 8 C, D and F). Overall, all RNA vaccines with mutation of glycosites elicited antibody

and CD4⁺ and CD8⁺ T cell responses and relative cytokines, with a stronger IFN γ -producing CD8⁺ T cell response in mice immunized with S-(deg-S2) and S-(S2-1194). These results suggested that glycosylation on S2 regulated the T cell response and expression of cytokines.

EXAMPLE 6: Deglycosylation on S2 induced unfolded protein response

[00208] To investigate how glycosylation on S2 affected immune response, HEK293 cells were transfected with the prefusion stabilized S protein expression plasmid of variants. It was shown that S-(deg-S2) and S-(S2-1194) did not express well, but the levels of S-(deg-S2) and S-(S2-1194) proteins were restored to some extent after treatment with MG132, a proteasome inhibitor (Fig. 9). In vitro translation assay showed that mutation of glycosites in the mRNA sequence did not affect the efficiency of translation (Fig. 10). These results suggested that removal of glycosylation from S2 caused degradation of translated protein *in vivo*. To study whether removal of glycosylation on S2 led to the expression of misfolded or unfolded protein, the plasma membrane and endoplasmic reticulum (ER) were isolated to examine the distribution of S protein, as misfolded or unfolded S proteins could be stacked in ER for refolding or destroyed through ER-associated degradation. After HEK293 cells were transfected with the mRNA at 48 hrs, all variant S proteins existed in plasma membrane, cytosol and ER. However, there was more protein of WT and S-(deg-RBD) in the plasma membrane, whereas there was more protein of S-(deg-S2) and S-(S2-1194) in the ER (Fig. 11). In addition, HEK293 cells were transfected with the mRNA that encoded the soluble pre-fusion version of S-(deg-S2) and S-(S2-1194) protein could not be secreted to the medium (Fig. 16), suggesting that deglycosylation in S2 affected the folding of S protein. Since the increased unfolded S protein in the ER would trigger the unfolded protein response (UPR), the UPR marker proteins BiP/GRP78, XBP1 and p-eIF2 α were examined in RNA transfected HEK293 cells at 48 hrs. The results showed that BiP/GRP78 and XBP1 were upregulated, and the level of p-eIF2 α was stronger than the WT and the S-(deg-RBD) groups in S-(deg-S2) and S-(S2-1194) transfected cells (Fig. 12). In addition, removal of glycosylation on S2 induced more apoptosis cells than WT and S-(deg-RBD) did (Fig. 13), suggesting that deglycosylation of S2 induced a higher level of ER stress than the WT and S-(deg-RBD) protein, and glycosylation on S2 affected protein folding to regulate the expression of UPR.

EXAMPLE 7: Glycosylation on S2 affected MHC I expression on dendritic cells (DCs).

[00209] To study whether UPR leads to the biased immune response, the major histocompatibility complex class I (MHC I) and class II (MHC II) on DCs, which are essential for presentation of the internalized molecules after processing, were measured by flow cytometry. After DCs incubated with variants of mRNA vaccine, MHC I / II were upregulated among all

vaccines, and the mRNA vaccine of S-(deg-S2) or S-(S2-1194) induced more MHC I expression DCs than WT and S-(deg-RBD) did (Fig. 14 and Fig. 17), suggesting that UPR regulated the expression of MHC I on DCs.

EXAMPLE 8: The glycosites in spike protein affect the stability of S protein and subsequently the CD8⁺ T cell response.

[00210] To study which glycosites regulated the host immune response, we used S-(deg-RBD) vaccine as the model system because it induced similar level of antibody as WT and had better neutralization activity against the four variants of concern than WT. Here, we removed the RBD glycosites and glycosite N-801 (S-(deg-RBD-801)) or glycosite N-1194 (S-(deg-RBD-1194)) in S2 as these glycosites involved the S protein expression, especially the glycosite N-1194 that involved in the integrity of S protein and its binding affinity. Since the glycosite N-122, N-165 and N-234 regulated the structure of RBD and affected the neutralization activity of antibody, we removed these glycosites to form the S-(deg-RBD-122-165-234) vaccine (Fig. 15). After HEK293 cells were transfected with the prefusion stabilized S protein expression plasmid of variants, it was shown that S-(deg-RBD-801) and S-(deg-RBD-122-165-234) did not express well and S-(deg-RBD-1194) dramatic decreased the protein expression, but the levels of these variant of proteins were restored to some extent after treatment with MG132, suggested that variant of vaccine induced the degradation of translated protein (Fig. 21A). After confirming the expression of the variant prefusion S protein form mRNA-LNP in HEK293 cell line (Fig. 21B), mice were immunized with varies vaccine. It showed that S-(deg-RBD-801), S-(deg-RBD-1194) and S-(deg-RBD-122-165-234) induced less IgG titer against fully glycosylated WT S (Fig. 18 A) and RBD protein (Fig. 18 B), but S-(deg-RBD-801) and S-(deg-RBD-122-165-234) had higher IgG titer against de-glycosylated RBD antigen (Fig. 18 C) in ELISA assay as compared to the unmodified mRNA, suggesting that glycosylation on these glycosites regulated the antibody production. To analyze the neutralization activity of antibodies generated from immunized mice, the pseudovirus neutralization assay showed that the S-(deg-RBD-801), S-(deg-RBD-1194) and S-(deg-RBD-122-165-234) mRNA vaccines generated antibodies with reduced neutralization activity against WT pseudovirus (Fig. 19), but with better neutralization activity against the four variants of concern than WT (Fig. 19 and Table 2).

Table 2: The IG₅₀ for variant pseudovirus neutralization assay.

IC ₅₀ serum dilution	WT	B.1.1.7	B.1.351	P1	B.1.617.2
WT	7430.8	1024.5	680.4	790.4	880.4
S-(deg-RBD-801)	5020.6	2680.4	1098.5	1640.5	1850.5

S-(deg-RBD-1194)	3480.5	1120.5	780.5	910.6	1040.3
S-(deg-RBD-122-165-234)	5240.8	3168.8	1820.8	2140.5	2040.5

[00211] To characterize the T cell response, the splenocytes of immunized mice were incubated with the peptide pool of S, RBD and S2 protein, then the granzyme B (GrzB)-secreting T cells were measured by elispot analysis. It was shown that S-(deg-RBD-801), S-(deg-RBD-1194) and S-(deg-RBD-122-165-234) induced more GrzB-secreting T cells than WT did in all peptide pools, especially in S-(deg-RBD-1194) did (Fig. 20 and Fig. 22A, 22B). After the isolated CD4⁺ and CD8⁺ T cells were incubated with bone-marrow-derived DCs and WT S peptide pool to measure the IFN γ -secreting T cells by flow cytometry, there was no significant difference in the number of the IFN γ -secreting CD4⁺ T cells among all vaccinated mice (Fig. 22C), but S-(deg-RBD-801), S-(deg-RBD-1194) and S-(deg-RBD-122-165-234) mRNA vaccine induced more IFN γ -secreting CD8⁺ T cells, suggesting that these vaccine induced stronger CD8⁺ T cell response (Fig. 20). Overall, S-(deg-RBD-801), S-(deg-RBD-1194) and S-(deg-RBD-122-165-234) decreased the level of the protein expression in plasmid system and induced less antibody, but increased the CD8⁺ T cell response, especially in S-(deg-RBD-1194) did. These results suggested that the glycosylation on the glycosites that involved in the folding or stability of S regulated the host immune response.

EXAMPLE 9: Design and Synthesis of guanidine-based poly(disulfide)s.

[00212] We designed a series of polymers and guanidine-based and/or zwitterionic head groups attached to a lipid tail and explored their ability to deliver spike mRNA. As shown in Fig. 23, propagator P1 containing guanidine groups and the multivalent display propagator P2, facilitate the adherence of mRNA with polymers by forming strong salt bridges between guanidiniums and the phosphates in mRNA. Once the polymer encapsulated mRNA reached the cytoplasm, the disulfide linkers could be degraded by intracellular glutathione to release mRNA. In addition, the degraded disulfide monomers also decrease the cytotoxicity by avoiding the accumulation of high molecular weight polymers inside the cell.

[00213] To prepare polymers, the mono guanidine containing disulfide monomer were synthesized according to previous reported procedures (Gasparini, G.; Bang, E.-K.; Molinard, G.; Tulumello, D. V.; Ward, S.; Kelley, S. O.; Roux, A.; Sakai, N.; Matile, S., *J. Am. Chem. Soc.* 2014, 136, 6069-6074), and the tri-guanidine disulfide monomer was synthesized from nitrilotriacetic acid linker to present trimeric guanidine monomer. The propagator **Pb**, containing

two strained disulfides beside the guanidine group was designed to form a branched configuration of polymer. Propagators **P3** and **Pz** were designed as a spacer which may facilitate the entrapped molecule to escape from endosome.

[00214] Fig. 24 depicts designed structure of initiators (I0), propagators (P1, P2), and the polymerization/depolymerization process.

[00215] The polymerization of **P1**, **P2**, **P3** and **Pb** was conducted in degassed water solution at room temperature. In brief, in the presence of 5 mM initiator **1** and 200 mM propagator **P** in 1 M pH 7 TEOA buffer was vigorously stirred for 30 minutes. Termination was done by adding 0.5 M iodoacetamide. To screen the optimal polymer for efficient intracellular delivery of mRNA, co-polymerization of different propagators was conducted, and their encapsulation ability and transfection efficiency was evaluated by GFP encoding mRNA in HEK293T cells. Co-polymer (**P1/P3**) and (**P2/P3**) were prepared in 2:1 ratio, (**P1/Pb**) and (**P2/Pb**) was prepared in 4:1 ratio.

[00216] In order to screen the optimal polymer for best encapsulation and efficient intracellular delivery of mRNA, 8 types of synthesized homopolymers and hetero-copolymers were examined for their ability to encapsulate GFP mRNA. As shown in Fig. 25A, all the copolymers possessed guanidine groups was able to inhibit the migration of GFP mRNA with N/P ratio = 10. Particularly, **P2/P3** and **P2/Pb** copolymer with triguanidine moiety resulted in higher capability to complex with the mRNA, which was comparable with result mediated by traditional used transfection agent polyethyleneimine (PEI). The average size of the resulting nanocluster **P2/P3**-mRNA complex was ca. 90 nm by TEM (Fig. 25B). In addition, the branched polymers **P1/Pb** and **P2/Pb** also showed similar ability for encapsulation of mRNA. Branched guanidiniums have attracted attention due to their good accessibility and flexibility to functional materials.

[00217] Next, we evaluated the transfection efficiency of GFP-mRNA in HEK293T cells by using different copolymers. First, we found that P1/P3 copolymer exhibited good ability to transfect the mRNA at N/P = 10, which is more efficient than P1 or P3 alone and the PEI (Fig. 26A). The result indicated that either P1 or P3 polymer was not favorable for intracellular delivery, while the bifunctional copolymer P1/P3 greatly transfected the mRNA. The randomly spacing of the guanidine group by copolymerization with diethylenetriamine spacer may be suitable to fit the phosphate charge of nucleotide. We then studied the impact of different copolymers in transfection activity. In Fig. 26B, a combination of P2/P3 showed increased GFP expression, indicating that the triguanidine group enhanced the encapsulation of mRNA, and could also liberate the mRNA once delivered to cytoplasm. However, the branched poly(disulfide)s P1/Pb and P2/Pb did not

reveal a satisfactory release of mRNA cargoes, although complete complexation with mRNA was done by using P2/Pb (Fig. 26A). We further optimized the N/P ratio (1, 5, 10, and 20) utilizing the outstanding copolymer P2/P3. The best performance for the transfection is found to be 10 (Fig. 26C).

[00218] Based on the results of GFP mRNA, wild type spike mRNA was prepared and encapsulated by polyGu at different N/P ratios (Fig. 27). PolyGu showed good ability to neutralize the charge of spike mRNA and successfully captured it at the N/P ratio of 1.

[00219] Next, we transfected spike mRNA in HEK293T cells and performed western blot. HEK293T cells were transfected with 3 μ g spike mRNA. 48 hours post transfection, cells were analysed for spike expression via western blotting using spike-specific antibody. The result showed a significant band of SARS-Cov-2 spike at around 250 kDa and PBS buffer with spike mRNA was employed as negative control (Fig. 28, which proved the feasibility of polyGu as a nanocarrier for mRNA transfection in vitro. In addition, polyGu did not exhibit any apparent cytotoxicity up to 10 μ g (50 μ g/mL) (Fig. 29). Whereas the lipid nanoparticles (LNP) exhibited much higher toxicity to cells with high complexes loading.

[00220] In conclusion, we have developed a series of poly(disulfide)s and demonstrated that a combination of guanidyl group and zwitterionic spacer exhibited great efficiency for mRNA delivery in vitro. The efficient intracellular delivery through thiol-mediated uptake pathway by strained disulfides is cleavable under the intracellular glutathione. In addition, the degradation of polymers also minimizes the cytotoxicity as compared to the commonly used LNP. against SARS-Cov-2.

[00221] All publications and patent applications cited in this specification are herein incorporated by reference as if each individual publication or patent application were specifically and individually indicated to be incorporated by reference.

[00222] Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it will be readily apparent to those of ordinary skill in the art in light of the teachings of this invention that certain changes and modifications may be made thereto without departing from the spirit or scope of the appended claim.

CLAIMS

What is claimed is:

1. A modified nucleic acid molecule encoding a modified spike protein comprising one or more amino acid substitutions of asparagine (N) to glutamine (Q) at N-linked glycosylation sequons (N-X-S/T), wherein X is any amino acid residue except proline, and S/T denotes a serine or threonine residue.
2. The modified nucleic acid molecule of claim 1, which comprises one or more amino acid deletions or additions at N-linked glycosylation sequons (N-X-S/T) to eliminate N-linked glycan sequons.
3. The modified nucleic acid molecule of claim 1, which comprises one or more amino acid substitutions of S/T to alanine (A) at O-linked glycosylation sites to eliminate O-linked glycosylation sites.
4. The modified nucleic acid molecule, which is an mRNA or a double-strand or single-strand DNA.
5. The modified nucleic acid molecule of any of claims 1 to 4, wherein the modified spike protein is derived from a SARS-CoV-2 spike protein.
6. The modified nucleic acid molecule of any of claims 1 to 4, wherein the spike protein comprises an amino acid sequence of SEQ ID NO: 2, 16, 18 or 20, or amino acid sequence having at least about 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 2, 16, 18 or 20.
7. The modified nucleic acid molecule of claim 6, wherein the nucleic acid molecule encoding the amino acid sequence of SEQ ID NO: 2, 16, 18 or 20 is an mRNA comprising the nucleotide sequence of SEQ ID NO: 1, 15, 17 or 19 respectively, or a nucleotide sequence having at least about 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 1, 15, 17 or 19 respectively.
8. The modified nucleic acid molecule of any of claims 1 to 4, wherein the modified spike protein comprises an amino acid sequence of SEQ ID NO: 4, 22, 24 or 26, wherein the modified spike protein comprises a receptor binding domain (RBD) lacking glycosylation sites.

9. The modified nucleic acid molecule of claim 8, wherein the modified nucleic acid molecule encoding the amino acid sequence of SEQ ID NO: 4, 22, 24 or 26 comprises the nucleotide sequence of SEQ ID NO: 3, 21, 23 or 25 respectively, or a nucleotide sequence having at least about 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 3, 21, 23 or 25 respectively.
10. The modified nucleic acid molecule of any of claims 1 to 4, wherein the modified spike protein comprises an amino acid sequence of SEQ ID NO: 6, 28, 30 or 32, wherein the modified spike protein comprises a S2 subunit lacking glycosylation sites.
11. The modified nucleic acid molecule of claim 10, wherein the modified nucleic acid molecule encoding the amino acid sequence of SEQ ID NO: 6, 28, 30 or 32 comprises the nucleotide sequence of SEQ ID NO: 5, 27, 29 or 31 respectively, or a nucleotide sequence having at least about 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 5, 27, 29 or 31 respectively.
12. The modified nucleic acid molecule of any of claims 1 to 4, wherein the modified spike protein described herein comprises an amino acid sequence of SEQ ID NO: 8 or 34, wherein the modified spike protein comprises a S2 subunit that consists of a single glycosylation site. In some embodiments, the single glycosylation site is at the position N1194.
13. The modified nucleic acid molecule of claim 12, wherein the modified nucleic acid molecule encoding the amino acid sequence of SEQ ID NO: 8 or 34 comprises the nucleotide sequence of SEQ ID NO: 7 or 33 respectively, or a nucleotide sequence having at least about 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 7 or 33 respectively.
14. The modified nucleic acid molecule of any of claims 1 to 4, wherein the modified spike protein comprises an amino acid sequence of SEQ ID NO: 10 or 36, wherein the modified spike protein comprises a receptor binding domain (RBD) lacking glycosylation sites, and an amino acid substitution of N801 to Q801.
15. The modified nucleic acid molecule of claim 14, wherein the modified nucleic acid molecule encoding the amino acid sequence of SEQ ID NO: 10 or 36 comprises the nucleotide sequence of SEQ ID NO: 9 or 35 respectively, or a nucleotide sequence

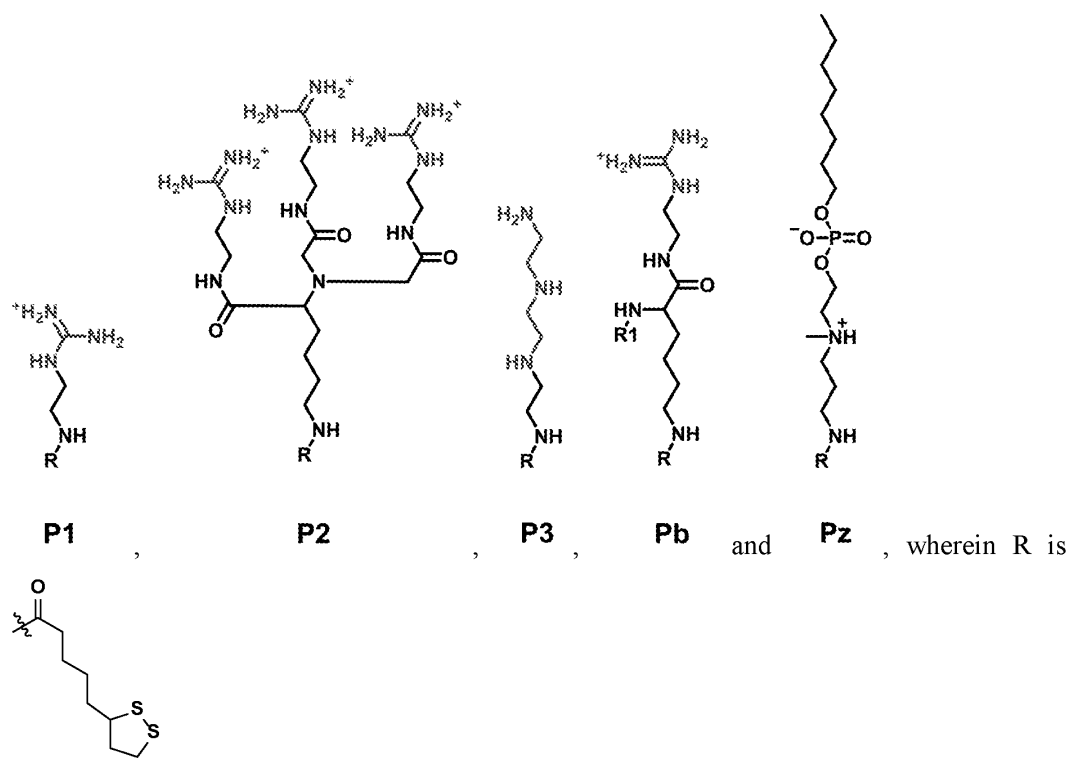
having at least about 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 9 or 35 respectively.

16. The modified nucleic acid molecule of any of claims 1 to 4, wherein the modified spike protein comprises an amino acid sequence of SEQ ID NO: 12 or 38, wherein the modified spike protein comprises a receptor binding domain (RBD) lacking glycosylation sites, and an amino acid substitution of N1194 to Q1194.
17. The modified nucleic acid molecule of claim 16, wherein the modified nucleic acid molecule encoding the amino acid sequence of SEQ ID NO: 12 or 38 comprises the nucleotide sequence of SEQ ID NO: 11 or 37 respectively, or a nucleotide sequence having at least about 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 11 or 37 respectively.
18. The modified nucleic acid molecule of any of claims 1 to 4, wherein the modified spike protein described herein comprises an amino acid sequence of SEQ ID NO: 14 or 40, wherein the modified spike protein comprises a modified receptor binding domain (RBD) lacking glycosylation sites, and amino acid substitutions of N122 to Q122, N165 to Q165, and N234 to Q234.
19. The modified nucleic acid molecule of claim 18, wherein the modified nucleic acid molecule encoding the amino acid sequence of SEQ ID NO: 14 or 40 comprises the nucleotide sequence of SEQ ID NO: 13 or 39 respectively, or a nucleotide sequence having at least about 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 13 or 39 respectively.
20. The modified nucleic acid molecule of any of claims 1 to 4, wherein the modified spike protein comprises a S1 subunit lacking glycosylation sites.
21. The modified nucleic acid molecule of any of claims 1 to 4, wherein the modified spike protein comprises both S1 and S2 subunits lacking glycosylation sites.
22. The modified nucleic acid molecule of any of claims 1 to 4, wherein the mRNA for S-(deg-RBD) has the sequence of SEQ ID NO: 3, 21, 23 or 25, the mRNA or DNA for S-(deg-S2) has the sequence of SEQ ID NO: 5, 27, 29 or 31, the mRNA or DNA for S-(deg-S2-1194) has the sequence of SEQ ID NO: 7 or 33, the mRNA or DNA for S-

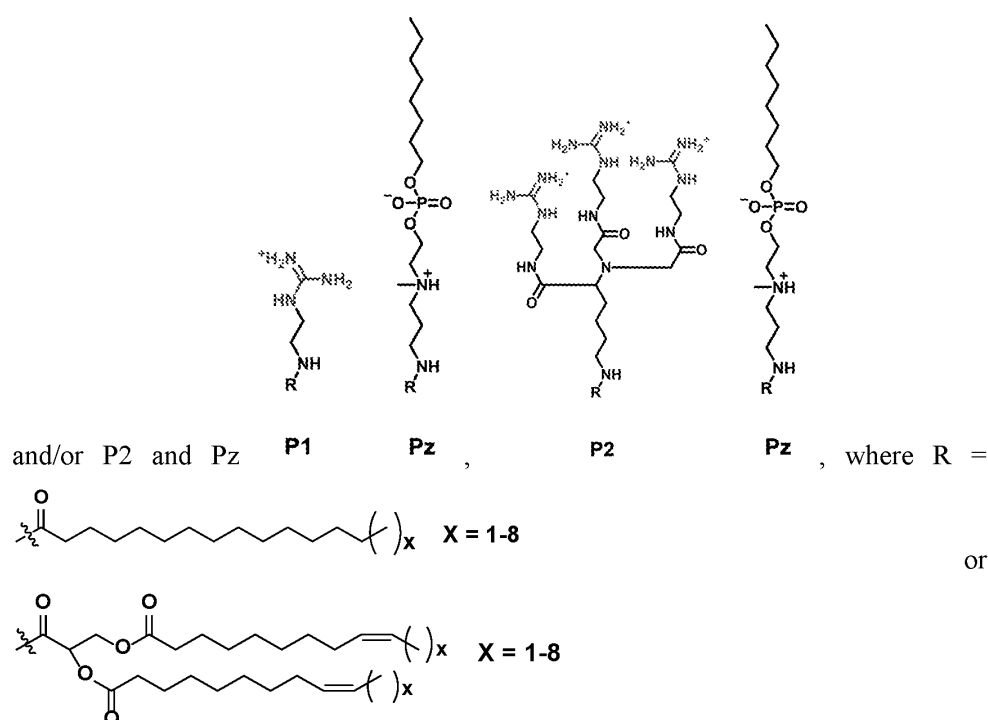
(deg-RBD-801) has the sequence of SEQ ID NO: 9 or 35, the mRNA or DNA for S-
(deg-RBD-1194) has the sequence of SEQ ID NO: 11 or 37, the mRNA or DNA for S-
(deg-RBD-122-165-234) has the sequence of SEQ ID NO: 13 or 39.

23. The modified nucleic acid molecule of any of claims 1 to 22, which can be used as a coronavirus mRNA vaccine.
24. The modified nucleic acid molecule of any of claims 1 to 23, wherein immunization of the modified nucleic acid molecule causes the upregulation of BiP/GRP78, XBP1 and p-eIF2 α to induce cell apoptosis and CD8 T-cell response.
25. The modified nucleic acid molecule of any of claims 1 to 23, wherein the coronavirus (CoV) is selected from the group consisting of alpha-CoV, beta-CoV, gamma-CoV, delta-CoV2, and omicron-CoV2.
26. A combo vaccine, comprising the mRNA vaccine of claim 23 and one or more additional vaccines.
27. The combo vaccine of claim 26, wherein the additional vaccine is selected from one or more COVID-19 vaccine, influenza (flu) vaccine, adenovirus vaccine, anthrax vaccine, cholera vaccine, diphtheria vaccine, hepatitis A or B vaccine, HPV vaccine, measles vaccine, mumps vaccine, smallpox vaccine, rotavirus vaccine, tuberculosis vaccine, pneumococcal vaccine and Haemophilus influenzae type b vaccine and any combination thereof.
28. A nanoparticle, which is guanidine-based and used as carrier for delivering the modified nucleic acid molecule of any one of claims 1 to 25 to a subject.
29. The nanoparticle of claim 28, which is a liposome or polymersome.
30. An mRNA nanocluster comprising the vaccine of claim 26 or 27 formulated in a lipid nanoparticle.
31. The mRNA nanocluster of claim 30, wherein the lipid nanoparticle is a biodegradable lipid nanoparticle.
32. The mRNA nanocluster of claim 30 or 31, wherein the lipid nanoparticle is guanidine-based polymers.

33. An mRNA nanocluster, comprising a lipid nanoparticle encapsulated with the mRNA vaccine, wherein the biodegradable lipid nanoparticle comprises guanidine-based and zwitterionic units, wherein the guanidine-based as well as zwitterionic groups attach to a lipid tail, and wherein the guanidine-based groups adhere to mRNA, thereby forming salt bridges between the guanidinium groups and the phosphates in the mRNA.
34. The nanoparticle of claim 28 or 29 or the mRNA nanocluster of any one of claims 30 to 33, wherein the guanidine-based polymers is P1, P2, P3, Pb or Pz:



35. The nanoparticle of claim 28 or 29 or the mRNA nanocluster of any one of claims 30 to 34, wherein the guanidine-based polymer forms a copolymer such as P1/P3 copolymer, P2/P3 copolymer, P1/Pb copolymer, P2/Pb copolymer, P1/Pz copolymer and P2/Pz copolymer.
36. The nanoparticle of claim 28 or 29 or the mRNA nanocluster of any one of claims 30 to 35, wherein the guanidine-based and zwitterionic nanoparticles are a mixture of P1,



37. The nanoparticle of claim 28 or 29 or the mRNA nanocluster of any one of claims 30 to 36, which has a nanoparticle/mRNA (N/P) ratio from about 1 to about 100.
38. The nanoparticle of claim 28 or 29 or the mRNA nanocluster of any one of claims 30 to 37, which has a nanoparticle/mRNA (N/P) ratio of about 10 or about 20.
39. A nanoparticle composition, comprising a nanoparticle attached with the mRNA vaccine of claim 23.
40. A method of preventing or treating a coronavirus infection, comprising administering an effective amount of the modified nucleic acid molecule of any one of claims 1 to 25, the nanoparticle of claim 28 or 29, the mRNA nanocluster of any one of claims 30 to 38 or a nanoparticle composition of claim 39 to a subject infected with, or at risk of being infected with, a coronavirus.
41. A method of boosting an adaptive immune response, comprising administering an effective amount of the modified nucleic acid molecule of any one of claims 1 to 25, the nanoparticle of claim 28 or 29, the mRNA nanocluster of any one of claims 30 to 38 or a nanoparticle composition of claim 39 to a subject.

42. The method of claim 41 or 42, wherein the modified nucleic acid molecule of any one of claims 1 to 25, the nanoparticle of claim 28 or 29, the mRNA nanocluster of any one of claims 30 to 38 or a nanoparticle composition of claim 39 is administered in an initial dose and two, three or four booster doses.
43. The method of claim 41 or 42, wherein the modified nucleic acid molecule of any one of claims 1 to 25, the nanoparticle of claim 28 or 29, the mRNA nanocluster of any one of claims 30 to 38 or a nanoparticle composition of claim 39 is administered in an initial dose and in at least one booster dose about one month, about two months, about three months, about four months, about five months, or about six months following the initial dose.
44. The method of claim 41 or 42, wherein the dose ranges from 5 μ g to 50 μ g of the modified nucleic acid molecule of any one of claims 1 to 25.
45. The method of claim 41 or 42, wherein the dose is 30 μ g.
46. The method of claim 41 or 42, wherein the modified nucleic acid molecule of any one of claims 1 to 25, the nanoparticle of claim 28 or 29, the mRNA nanocluster of any one of claims 30 to 38 or a nanoparticle composition of claim 39 is administered via intravenous route, intramuscular route, intradermal route, or subcutaneous route, or by infusion or nasal spray.
47. An immunogenic peptide, comprising at least one amino acid sequence selected from a group consisting of: TESIVRFPNITNL (SEQ ID NO: 41), NITNLCPFGEVFNATR (SEQ ID NO: 42), LYNSASFSTFK (SEQ ID NO: 43), LDSKVGGNYN (SEQ ID NO: 44), KSNLKPFERDIST (SEQ ID NO: 45), KPFERDISTEIYQAG (SEQ ID NO: 46), GPKKSTNLVKNNKC (SEQ ID NO: 47), NCDVVIGIVNNTVY (SEQ ID NO: 48), PELDSFKEELDKYFKNHTS (SEQ ID NO: 49), VNIQKEIDRLNEVA (SEQ ID NO: 50), NLNESLIDLQ (SEQ ID NO: 51) and LGKYEQYIKWP (SEQ ID NO: 52) or an amino acid sequence having at least about 99%, 98%, 97%, 96%, 95% or 90% identity to any of SEQ ID NOs: 41 to 52.
48. The immunogenic peptide of claim 47, comprising at least one amino acid sequence selected from a group consisting of SEQ ID NOs: 41 to 43 and 45 to 51.

49. An immunogenic composition, comprising an immunogenic peptide of claim 47 or 48.