SARS-COV-2 MRNA DOMAIN VACCINES

RELATED APPLICATIONS

This application claims the benefit under 35 U.S.C. § 119(e) of U.S. provisional application number 62/971,825, filed February 7, 2020, U.S. provisional application number 63/016,175, filed April 27, 2020, U.S. provisional application number 63/044,330, filed June 25, 2020, and U.S. provisional application number 63/063,137, filed August 7, 2020, each of which is incorporated by reference herein in its entirety.

BACKGROUND

Human coronaviruses are highly contagious enveloped, positive single-stranded RNA viruses of the *Coronaviridae* family. Two sub-families of *Coronaviridae* are known to cause human disease. The most important being the β-coronaviruses (betacoronaviruses). The β-coronaviruses are common etiological agents of mild to moderate upper respiratory tract infections. Outbreaks of novel coronavirus infections such as the infections caused by a coronavirus initially identified from the Chinese city of Wuhan in December 2019, however, have been associated with a high mortality rate death toll. This recently identified coronavirus, referred to as Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) (formerly referred to as a "2019 novel coronavirus," or a "2019-nCoV") has rapidly infected hundreds of thousands of people. The pandemic disease that the SARS-CoV-2 virus causes has been named by World Health Organization (WHO) as COVID-19 (Coronavirus Disease 2019). The first genome sequence of a SARS-CoV-2 isolate (Wuhan-Hu-1) was released by investigators from the Chinese CDC in Beijing on January 10, 2020 at Virological, a UK-based discussion forum for analysis and interpretation of virus molecular evolution and epidemiology. The sequence was then deposited in GenBank on January 12, 2020, having Genbank Accession number MN908947.1.

Currently, there is no specific treatment for COVID-19 or vaccine for SARS-CoV-2 infection. The continuing health problems and mortality associated with coronavirus infections, particularly the SARS-CoV-2 pandemic, are of tremendous concern internationally. The public health crisis caused by SARS-CoV-2 reinforces the importance of rapidly developing effective and safe vaccine candidates against these viruses.

SUMMARY

Provided herein, in some embodiments, are compositions (e.g., vaccines) that comprise one or more messenger ribonucleic acid (mRNA) molecules that encode(s) highly immunogenic antigen(s) capable of eliciting potent neutralizing antibody responses against SARS-CoV-2 antigens. The mRNA molecules described herein are used to express key neutralizing domains of the SARS-CoV-2 coronavirus spike (S) protein that are efficient at inducing protective immunity when used individually or in combination as an immunogenic composition or vaccine to protect people from infection by the natural virus and/or to reduce symptoms if infected.

The envelope S proteins of known betacoronaviruses determine the virus host tropism and entry into host cells and are critical for SARS-CoV-2 infection. The organization of the S protein is similar among betacoronaviruses, such as SARS-CoV-2, SARS-CoV, MERS-CoV, HKU1-CoV, MHV-CoV and NL63-CoV, including two subunits, S1 and S2, which mediate attachment and membrane fusion, respectively. The S1 subunit includes an N terminal domain (NTD) and a receptor binding domain (RBD).

The expression of subunit antigens focuses the immune response to specific subunits with minimal stimulation of memory B and T cells specific to other domains of the antigen that are shared with other related viruses. Data provided herein demonstrates that administration of mRNA encoding membrane bound or soluble SARS-CoV-2 S1 subunit antigen generated antibody titers to each of SARS-CoV-2 RBD antigen, NTD antigen, wildtype full-length S protein, and S protein having double proline mutations to stabilize the prefusion conformation. As shown herein, at all doses tested, a two-dose regimen (i.e., including a booster dose) was effective at inducing antibodies that could recognize and bind to SARS-CoV-2 WT S protein. Surprisingly, the induced titers were highest when measured against the double proline stabilized version of the S protein even though the double proline mutation is not found in the S1 subunit (the double proline mutation occurs in S2, and S2 was not present in the immunogen tested).

Additionally, both the NTD and RBD are known to be sites for binding of antibodies that neutralize virus activity. RBD in the case of SARS-CoV-2 is the receptor binding site of the spike protein which binds the angiotensin-converting enzyme 2 (ACE2). The NTD, the function of which is not thoroughly understood, seems to have a role in binding sugar moieties and in facilitating the conformational transition of the spike protein from prefusion to a post fusion conformation. Regardless, both the NTD and RBD domains induce high binding antibody and neutralizing antibody titers as shown herein.

For example, quite surprisingly, the data provided in some embodiments herein show that while sera from the administration of mRNA encoding a membrane bound RBD antigen (RBD-

TM) or a membrane bound NTD antigen (NTD-TM) showed immunogenicity to the SARS-CoV-2 S1/S2 spike protein, the 50:50 combination of the two mRNAs (and thus the two antigens) generated unexpectedly high, synergistic, neutralizing antibody titers to the SARS-CoV-2 S1/S2 spike protein.

Thus, some aspects of the present disclosure provide compositions comprising an mRNA encoding a functional domain of a SARS-CoV-2 S protein capable of inducing an immune response, such as a neutralizing antibody response, to a SARS-CoV-2. In some embodiments, the mRNA is formulated in a lipid nanoparticle.

In some aspects an mRNA comprising an open reading frame (ORF) that encodes at least two domains of a SARS-CoV-2 Spike protein, and less than the full length spike protein is provided. A spike protein that is less than the full length spike protein is one or more domains and/or subunits of the spike protein having at least one amino acid less than the full length spike protein or a fusion protein having one or more domains linked together in an non-natural order or sequence. In some embodiments one of the two domains is an N-terminal domain (NTD) of a SARS-CoV-2 Spike protein. In some embodiments one of the two domains is a receptor binding domain (RBD) of a SARS-CoV-2 Spike protein. In some embodiments the ORF encodes a transmembrane domain (TD) linked to the NTD and/or RBD. In some embodiments the TD is an influenza hemagglutinin transmembrane domain. In some embodiments the ORF comprises NTD - RBD - TM. In some embodiments the at least two domains are linked through a cleavable or non-cleavable linker. In some embodiments the non-cleavable linker is a glycine-serine (GS) linker. In some embodiments the GS linker 4-15 amino acids. In some embodiments the linker is a pan HLA DR-binding epitope (PADRE). In some embodiments the ORF encodes a signal peptide. In some embodiments the signal peptide is linked to the NTD. In some embodiments the signal peptide is linked to the RBD. In some embodiments the signal peptide is heterologous to SARS-CoV-2. In some embodiments the at least two domains are soluble. In some embodiments the ORF encodes a trafficking signal domain. In some embodiments the trafficking signal domain is a macrophage marker. In some embodiments the macrophage marker CD86 and/or CD11b. In some embodiments the trafficking signal domain is a VSV-G cytosolic tail (VSVGct). In some embodiments one of the two domains is a first repetitive heptapeptide: HPPHCPC (HR1) of a SARS-CoV-2 Spike protein. In some embodiments one of the two domains is a second repetitive heptapeptide: HPPHCPC (HR2) of a SARS-CoV-2 Spike protein. In some embodiments the ORF encodes a transmembrane domain (TD) linked to the HR1 and/or HR2. In some embodiments the TD is an influenza hemagglutinin transmembrane domain. In some

embodiments the ORF encodes a fusion peptide (FP). In some embodiments the ORF encodes a CT tail.

In some aspects an mRNA comprising an open reading frame (ORF) that encodes a receptor binding domain (RBD) of a SARS-CoV-2 Spike protein is provided. In some embodiments the RBD is soluble. In some embodiments the RBD is linked to a transmembrane domain, optionally an influenza hemagglutinin transmembrane domain.

The details of one or more embodiments of the invention are set forth in the description below. Other features or advantages of the present invention will be apparent from the following drawings and detailed description of several embodiments, and also from the appended claims.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 Schematic representation of wild-type and 2P spike protein antigens encoded by mRNAs of the invention; signal peptide (SP), no fill, N-terminal domain (NTD), dotted; receptor-binding domain (RBD), downward diagonal stripes; subdomain 1 (SD1), horizontal stripes; subdomain 2 (SD2), wave; fusion peptide (FP), upward diagonal stripes; heptad repeat 1 (HR1) weave; heptad repeat 2 (HR2) diagonal brick; (TM), vertical stripes; and cytoplasmic tail (CT), brick.

FIG. 2 shows examplary linear designs of the antigens encoded by the mRNAs described in Examples 1-3.

FIG. 3 shows sequence alignments of the antigens depicted in FIG. 2.

FIG. 4 shows examplary linear designs of the antigens encoded by the mRNAs described in Examples 4-6.

FIG. 5 shows sequence alignments of various S1 subunit antigens described herein.

FIG. 6 shows examplary linear designs of the antigens encoded by the mRNAs described in Examples 7 and 8.

FIG. 7 shows correlations of neutralization and ELISA titers.

FIGs. 8A-8C show serum IgG1 and IgG2a Titers at Day 36 following a Day 1 prime and Day 21 boost dose in mice with mRNA encoding NTD-RBD-TM in an LNP.

DETAILED DESCRIPTION

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is a newly emerging respiratory virus with high morbidity and mortality. SARS-CoV-2 has rapidly spread around the world compared with SARS-CoV, which appeared in 2002, and Middle East respiratory syndrome coronavirus (MERS-CoV), which emerged in 2012. The World Health Organization

(WHO) reports that, as of July 6, 2020, the current outbreak of COVID-19 has almost 11.5 million confirmed cases worldwide with more than 530,000 deaths. New cases of COVID-19 infection are on the rise and are still increasing rapidly. It is thus crucial that a variety of safe and effective vaccines and drugs be developed to prevent and treat COVID-19 and reduce the serious impact that COVID-19 is having across the world. Vaccines and drugs made using a variety of modalities, and vaccines having improved safety and efficacy, are imperative. Their remains a need to accelerate the advanced design and development of vaccines and therapeutic drugs against coronavirus disease 2019 (COVID-19).

On January 7, 2020, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) was identified as the etiological agent of a novel pneumonia that emerged in December 2019, in Wuhan City, Hubei province in China (Lu H. et al. (2020) J Med Virol. Apr; 92(4):401-402.). Soon after, the virus caused an outbreak in China and has spread to the world. According to the analysis of genomic structure of SARS-CoV-2, it belongs to β-coronaviruses (CoVs) (Chan et al. 2020 Emerg Microbes Infect.; 9(1):221-236).

A key protein on the surface of coronavirus is the Spike protein. A large variety of mRNA constructs have been designed and are disclosed herein. When formulated in appropriate delivery vehicles mRNA encoding Spike antigen, subunits and domains thereof are capable of inducing a strong immune response against SARS-CoV-2, thus producing effective and potent mRNA vaccines. Administration of the mRNA encoding various Spike protein antigens, in particular, Spike protein subunit and domain antigens, results in delivery of the mRNA to immune tissues and cells of the immune system where it is rapidly translated into proteins antigens. Other immune cells, for example, B cells and T cells, are then able to recognize and mount and immune response develop an immune response against the encoded protein and ultimately create a long-lasting protective response against the coronavirus. Low immunogenicity, a drawback in protein vaccine development due to poor presentation to the immune system or incorrect folding of the antigens, is avoided through the use of the highly effective mRNA vaccines encoding spike protein, subunits and domains thereof disclosed herein.

The present disclosure provides compositions (e.g., mRNA vaccines) that elicit potent neutralizing antibodies against coronavirus antigens. In some embodiments, a composition includes mRNA encoding at least one (e.g., one, two, or more) coronavirus antigen, such as a SARS-CoV-2 antigen. In some embodiments, the mRNA encodes a spike protein domain, such as a receptor binding domain (RBD), an N-terminal domain (NTD), or a combination of an RBD and NTD.

Some aspects of the present disclosure provide a messenger ribonucleic acid (mRNA) comprising an open reading frame encoding a fusion protein comprising a receptor binding domain (RBD) of a SARS-CoV-2 Spike protein and a protein transmembrane domain, e.g., a naturally occurring or heterologous transmembrane domain.

In some embodiments, the protein transmembrane domain is an influenza hemagglutinin transmembrane domain.

In some embodiments, the fusion protein comprises an amino acid sequence having at least 80% identity to the amino acid sequence of SEQ ID NO: 77.

In some embodiments, the fusion protein comprises an amino acid sequence having at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 77.

In some embodiments, the fusion protein comprises the amino acid sequence of SEQ ID NO: 77.

In some embodiments, the open reading frame comprises a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 76.

In some embodiments, the wherein the open reading frame comprises a nucleotide sequence having at least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 76.

In some embodiments, the open reading frame comprises the nucleotide sequence of SEQ ID NO: 76.

Other aspects of the present disclosure provide a messenger ribonucleic acid (mRNA) comprising an open reading frame encoding a fusion protein comprising an amino (N)-terminal domain of a SARS-CoV-2 Spike protein and a transmembrane domain.

In some embodiments, the transmembrane domain is an influenza hemagglutinin transmembrane domain.

In some embodiments, the fusion protein comprises an amino acid sequence having at least 80% identity to the amino acid sequence of SEQ ID NO: 47.

In some embodiments, the fusion protein comprises an amino acid sequence having at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 47.

In some embodiments, the fusion protein comprises the amino acid sequence of SEQ ID NO: 47.

In some embodiments, the open reading frame comprises a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 46.

In some embodiments, the open reading frame comprises a nucleotide sequence having at least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 46.

In some embodiments, the open reading frame comprises the nucleotide sequence of SEQ ID NO: 46.

Yet other aspects of the present disclosure provide a messenger ribonucleic acid (mRNA) comprising an open reading frame encoding a fusion protein comprising a receptor binding domain of a SARS-CoV-2 Spike protein linked to an amino (N)-terminal domain of a SARS-CoV-2 Spike protein, optionally via a linker.

In some embodiments, the fusion protein further comprises a transmembrane domain.

In some embodiments, the fusion protein comprises an amino acid sequence having at least 80% identity to the amino acid sequence of SEQ ID NO: 92.

In some embodiments, the fusion protein comprises an amino acid sequence having at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 92.

In some embodiments, the fusion protein comprises the amino acid sequence of SEQ ID NO: 92.

In some embodiments, the open reading frame comprises a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 91.

In some embodiments, the open reading frame comprises a nucleotide sequence having at least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 91.

In some embodiments, the open reading frame comprises the nucleotide sequence of SEQ ID NO: 91.

In some embodiments, the mRNA further comprises a 5′ untranslated region (UTR), optionally comprising the nucleotide sequence of SEQ ID NO: 131 or 2.

In some embodiments, the mRNA further comprises a 3′ untranslated region (UTR), optionally comprising the nucleotide sequence of SEQ ID NO: 132 or 4.

In some embodiments, the mRNA further comprises a 5′ cap, optionally 7mG(5')ppp(5')NlmpNp.

In some embodiments, the mRNA further comprises a polyA tail, optionally having a length of about 100 nucleotides.

In some embodiments, the mRNA comprises a chemical modification, optionally 1-methylpseudouridine.

Some aspects of the present disclosure provide a composition comprising the mRNA of any one of the preceding paragraphs.

Other aspects of the present disclosure provide a composition comprising at least two of the mRNA of any one of the preceding paragraphs.

Other aspects of the present disclosure provide a composition comprising: (a) a messenger ribonucleic acid (mRNA) comprising an open reading frame encoding a fusion protein comprising a receptor binding domain (RBD) of a SARS-CoV-2 Spike protein and a protein transmembrane domain; and (b) an mRNA comprising an open reading frame encoding a fusion protein comprising an amino (N)-terminal domain of a SARS-CoV-2 Spike protein and a transmembrane domain. In some embodiments, the ratio of the mRNA of (a) to the mRNA of (b) is about 1:1, e.g., 1:2, 1:3, 21:, or 3:1. In some embodiments, at least 50% of the mRNA of a composition is the mRNA of (a). For example, at least 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the mRNA of a composition is the mRNA of (a). In some embodiments, at least 50% of the mRNA of a composition is the mRNA of (b). For example, at least 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% of the mRNA of a composition is the mRNA of (b).

In some embodiments, the protein transmembrane domain is an influenza hemagglutinin transmembrane domain.

In some embodiments, the fusion protein of (a) comprises an amino acid sequence having at least 80% identity to the amino acid sequence of SEQ ID NO: 77.

In some embodiments, the fusion protein of (a) comprises an amino acid sequence having at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 77.

In some embodiments, the fusion protein of (a) comprises the amino acid sequence of SEQ ID NO: 77.

In some embodiments, the open reading frame of (a) comprises a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 76.

In some embodiments, the open reading frame of (a) comprises a nucleotide sequence having at least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 76.

In some embodiments, the open reading frame of (a) comprises the nucleotide sequence of SEQ ID NO: 76.

In some embodiments, the fusion protein of (b) comprises an amino acid sequence having at least 80% identity to the amino acid sequence of SEQ ID NO: 47.

In some embodiments, the fusion protein of (b) comprises an amino acid sequence having at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 47.

In some embodiments, the fusion protein of (b) comprises the amino acid sequence of SEQ ID NO: 47.

In some embodiments, the open reading frame of (b) comprises a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 46.

In some embodiments, the open reading frame of (b) comprises a nucleotide sequence having at least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 46.

In some embodiments, the open reading frame of (b) comprises the nucleotide sequence of SEQ ID NO: 46.

In some embodiments, the mRNA is formulated in a lipid nanoparticle.

In some embodiments, the composition further comprises a lipid nanoparticle.

In some embodiments, the mRNA of (a) is formulated in a lipid nanoparticle, and the mRNA of (b) is formulated in a lipid nanoparticle.

In some embodiments, the lipid nanoparticle comprises a cationic lipid.

In some embodiments, the lipid nanoparticle further comprises a neutral lipid.

In some embodiments, the lipid nanoparticle further comprises a sterol.

In some embodiments, the lipid nanoparticle further comprises a polyethylene glycol (PEG)-modified lipid.

In some embodiments, the lipid nanoparticle comprises an ionizable cationic lipid, a neutral lipid, a sterol, and a PEG-modified lipid.

In some embodiments, the ionizable cationic lipid is heptadecan-9-yl 8 ((2 hydroxyethyl)(6 oxo 6-(undecyloxy)hexyl)amino)octanoate (Compound 1).

In some embodiments, the neutral lipid is 1,2 distearoyl-sn-glycero-3-phosphocholine (DSPC).

In some embodiments, the sterol is cholesterol.

In some embodiments, the PEG-modified lipid is 1,2 dimyristoyl-sn-glycerol, methoxypolyethyleneglycol (PEG2000 DMG).

In some embodiments, the lipid nanoparticle comprises 20-60 mol% ionizable cationic lipid, 5-25 mol% neutral lipid, 25-55 mol% sterol, and 0.5-15 mol% PEG-modified lipid.

In some embodiments, the lipid nanoparticle comprises: 47 mol% ionizable cationic lipid; 11.5 mol% neutral lipid; 38.5 mol% sterol; and 3.0 mol% PEG-modified lipid; 48 mol%

ionizable cationic lipid; 11 mol% neutral lipid; 38.5 mol% sterol; and 2.5 mol% PEG-modified lipid; 49 mol% ionizable cationic lipid; 10.5 mol% neutral lipid; 38.5 mol% sterol; and 2.0 mol% PEG-modified lipid; 50 mol% ionizable cationic lipid; 10 mol% neutral lipid; 38.5 mol% sterol; and 1.5 mol% PEG-modified lipid; or 51 mol% ionizable cationic lipid; 9.5 mol% neutral lipid; 38.5 mol% sterol; and 1.0 mol% PEG-modified lipid.

In some embodiments, the lipid nanoparticle comprises: 47 mol% Compound 1; 11.5 mol% DSPC; 38.5 mol% cholesterol; and 3.0 mol% PEG2000 DMG; 48 mol% Compound 1; 11 mol% DSPC; 38.5 mol% cholesterol; and 2.5 mol% PEG2000 DMG; 49 mol% Compound 1; 10.5 mol% DSPC; 38.5 mol% cholesterol; and 2.0 mol% PEG2000 DMG; 50 mol% Compound 1; 10 mol% DSPC; 38.5 mol% cholesterol; and 1.5 mol% PEG2000 DMG; or 51 mol% Compound 1; 9.5 mol% DSPC; 38.5 mol% cholesterol; and 1.0 mol% PEG2000 DMG.

Further aspects of the present disclosure provide a method comprising administering to a subject the mRNA or the composition of any one of the preceding claims in an amount effective to induce in the subject a neutralizing antibody response against SARS-CoV-2.

Other aspects of the present disclosure provide a method comprising administering to a subject the mRNA or the composition of any one of the preceding claims in an amount effective to induce in the subject and a T cell immune response against SARS-CoV-2.

Some aspects of the present disclosure provide a messenger ribonucleic acid (mRNA) comprising an open reading frame (ORF) that encodes a coronavirus antigen capable of inducing an immune response, such as a neutralizing antibody response, to a SARS-CoV-2, wherein the antigen comprises a protein fragment or a functional protein domain of a SARS-CoV-2, optionally wherein the RNA is formulated in a lipid nanoparticle.

In some embodiments, the antigen is a functional protein domain.

In some embodiments, the protein domain is an N-terminal domain (NTD) of a SARS-CoV-2 Spike protein.

In some embodiments, the NTD is linked to a transmembrane domain, optionally an influenza hemagglutinin transmembrane domain.

In some embodiments, the antigen comprises an amino acid sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 47, optionally wherein the antigen comprises the amino acid sequence of SEQ ID NO: 47.

In some embodiments, the open reading frame comprises a nucleotide sequence having at least 70%, least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least

98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 46, optionally wherein the open reading frame comprises the nucleotide sequence of SEQ ID NO: 46.

In some embodiments, the protein domain is a receptor binding domain (RBD) of a SARS-CoV-2 Spike protein.

In some embodiments, the RBD is soluble.

In some embodiments, the antigen comprises an amino acid sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 62, optionally wherein the antigen comprises the amino acid sequence of SEQ ID NO: 62.

In some embodiments, the open reading frame comprises a nucleotide sequence having at least 70%, least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 61, optionally wherein the open reading frame comprises the nucleotide sequence of SEQ ID NO: 61.

In some embodiments, the RBD is linked to a transmembrane domain, optionally an influenza hemagglutinin transmembrane domain.

In some embodiments, the antigen comprises an amino acid sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 77, optionally wherein the antigen comprises the amino acid sequence of SEQ ID NO: 77.

In some embodiments, the open reading frame comprises a nucleotide sequence having at least 70%, least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 76, optionally wherein the open reading frame comprises the nucleotide sequence of SEQ ID NOs: 76.

In some embodiments, the NTD is linked to an RBD of a SARS-CoV-2 Spike protein to form an NTD-RBD fusion protein.

In some embodiments, the NTD-RBD fusion is linked to a transmembrane domain (TM), optionally an influenza hemagglutinin transmembrane domain, to form an NTD-RBD-TM protein.

In some embodiments, the antigen comprises an amino acid sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 92, optionally wherein the antigen comprises the amino acid sequence of SEQ ID NO: 92.

In some embodiments, the open reading frame comprises a nucleotide sequence having at least 70%, least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least

98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 91, optionally wherein the open reading frame comprises the nucleotide sequence of SEQ ID NO: 91.

In some embodiments, the NTD-RBD fusion comprises a C-terminal truncation.

In some embodiments, the antigen comprises an amino acid sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 107, optionally wherein the antigen comprises the amino acid sequence of SEQ ID NO: 107.

In some embodiments, the open reading frame comprises a nucleotide sequence having at least 70%, least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 106, optionally wherein the open reading frame comprises the nucleotide sequence of SEQ ID NO: 106.

In some embodiments, the NTD and/or RBD includes an extended region.

In some embodiments, the antigen comprises an amino acid sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of any one of SEQ ID NOs: 59, 86, 89, 116, 119, or 122, optionally wherein the antigen comprises the amino acid sequence of any one of SEQ ID NOs: 59, 86, 89, 116, 119, or 122.

In some embodiments, the open reading frame comprises a nucleotide sequence having at least 70%, least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of any one of SEQ ID NOs: 58, 85, 88, 115, 118, or 121, optionally wherein the open reading frame comprises the nucleotide sequence of any one of SEQ ID NOs: 58, 85, 88, 115, 118, or 121.

In some embodiments, the protein domain is an S1 subunit domain of a SARS-CoV-2 Spike protein.

In some embodiments, the S1 subunit is soluble.

In some embodiments, the antigen comprises an amino acid sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 5, optionally wherein the antigen comprises the amino acid sequence of SEQ ID NO: 5.

In some embodiments, the open reading frame comprises a nucleotide sequence having at least 70%, least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 3, optionally wherein the open reading frame comprises the nucleotide sequence of SEQ ID NO: 3.

In some embodiments, the S1 subunit is linked to a transmembrane domain, optionally an influenza hemagglutinin transmembrane domain.

In some embodiments, the antigen comprises an amino acid sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 17, optionally wherein the antigen comprises the amino acid sequence of SEQ ID NO: 17.

In some embodiments, the open reading frame comprises a nucleotide sequence having at least 70%, least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 16, optionally wherein the open reading frame comprises the nucleotide sequence of SEQ ID NO: 16.

In some embodiments, the S1 subunit has been modified to remove an RBD or a portion of an RBD of S protein.

In some embodiments, the antigen comprises an amino acid sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of any one of SEQ ID NOs: 20, 23, 26, 29, 32 or 35, optionally wherein the antigen comprises the amino acid sequence of any one of SEQ ID NOs: 20, 23, 26, 29, 32, or 35.

In some embodiments, the open reading frame comprises a nucleotide sequence having at least 70%, least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of any one of SEQ ID NOs: 19, 22, 25, 28, 41, or 34, optionally wherein the open reading frame comprises the nucleotide sequence of any one of SEQ ID NOs: 19, 22, 25, 28, 31, or 34.

In some embodiments, the S1 subunit is linked to an S2 subunit of an S protein.

In some embodiments, the S2 subunit is from a SARS-CoV-2 S protein.

In some embodiments, the S1 subunit is from an HKU1 S protein.

In some embodiments, the antigen comprises an amino acid sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 38, optionally wherein the antigen comprises the amino acid sequence of SEQ ID NO: 38.

In some embodiments, the open reading frame comprises a nucleotide sequence having at least 70%, least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 37, optionally wherein the open reading frame comprises the nucleotide sequence of SEQ ID NO: 37.

In some embodiments, the S1 subunit is from an OC43 S protein.

In some embodiments, the antigen comprises an amino acid sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 41, optionally wherein the antigen comprises the amino acid sequence of SEQ ID NO: 41.

In some embodiments, the open reading frame comprises a nucleotide sequence having at least 70%, least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 40, optionally wherein the open reading frame comprises the nucleotide sequence of SEQ ID NO: 40.

In some embodiments, the antigen further comprises a scaffold domain, optionally selected from ferritin, lumazine synthetase and a foldon.

In some embodiments, the scaffold domain is ferritin.

In some embodiments, the antigen comprises an amino acid sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 8 or 65, optionally wherein the antigen comprises the amino acid sequence of SEQ ID NO: 8 or 65.

In some embodiments, the open reading frame comprises a nucleotide sequence having at least 70%, least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 7 or 64, optionally wherein the open reading frame comprises the nucleotide sequence of SEQ ID NO: 7 or 64.

In some embodiments, the scaffold domain is lumazine synthetase.

In some embodiments, the antigen comprises an amino acid sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of any one of SEQ ID NOs: 11, 14, 68, or 71, optionally wherein the antigen comprises the amino acid sequence of any one of SEQ ID NOs: 11, 14, 68, or 71.

In some embodiments, the open reading frame comprises a nucleotide sequence having at least 70%, least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of any one of SEQ ID NOs: 10, 13, 67, or 70, optionally wherein the open reading frame comprises the nucleotide sequence of any one of SEQ ID NOs: 10, 13, 67, or 70.

In some embodiments, the scaffold domain is a foldon.

In some embodiments, the antigen comprises an amino acid sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of any one of SEQ ID NOs: 44, 50, 74, 80, 83, 101, 104

or 113, optionally wherein the antigen comprises the amino acid sequence of any one of SEQ ID NOs: 44, 50, 74, 80, 83, 101, 104 or 113.

In some embodiments, the open reading frame comprises a nucleotide sequence having at least 70%, least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of any one of SEQ ID NOs: 43, 49, 73, 79, 82, 100, 103, or 112, optionally wherein the open reading frame comprises the nucleotide sequence of any one of SEQ ID NOs: 43, 49, 73, 79, 82, 100, 103, or 112.

In some embodiments, the antigen further comprises a trafficking signal, optionally selected from macrophage markers, optionally CD86, CD11B and/or VSVGct.

In some embodiments, the antigen comprises an amino acid sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of any one of SEQ ID NOs: 95, 98, or 110, optionally wherein the antigen comprises the amino acid sequence of any one of SEQ ID NOs: 95, 98, or 110.

In some embodiments, the open reading frame comprises a nucleotide sequence having at least 70%, least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of any one of SEQ ID NOs: 94, 97, or 109, optionally wherein the open reading frame comprises the nucleotide sequence of any one of SEQ ID NOs: 94, 97, or 109.

In some embodiments, the mRNA is formulated in a lipid nanoparticle.

In some embodiments, the lipid nanoparticle comprises a cationic lipid, optionally an ionizable cationic lipid, a neutral lipid, a sterol, and/or a polyethylene glycol (PEG)-modified lipid. An ionizable cationic lipid is used interchangeably herein with ionizable lipid and cationic lipid to refer to an ionizable lipid. In some embodiments, the lipid nanoparticle comprises 40-50 mol% ionizable lipid, optionally 45-50 mol%, for example, 45-46 mol%, 46-47 mol%, 47-48 mol%, 48-49 mol%, or 49-50 mol% for example about 45 mol%, 45.5 mol%, 46 mol%, 46.5 mol%, 47 mol%, 47.5 mol%, 48 mol%, 48.5 mol%, 49 mol%, or 49.5 mol%. In some embodiments, the lipid nanoparticle comprises 30-45 mol% sterol, optionally 35-40 mol%, for example, 30-31 mol%, 31-32 mol%, 32-33 mol%, 33-34 mol%, 35-35 mol%, 35-36 mol%, 36-37 mol%, 38-38 mol%, 38-39 mol%, or 39-40 mol%. In some embodiments, the lipid nanoparticle comprises 5-15 mol% helper lipid, optionally 10-12 mol%, for example, 5-6 mol%, 6-7 mol%, 7-8 mol%, 8-9 mol%, 9-10 mol%, 10-11 mol%, 11-12 mol%, 12-13 mol%, 13-14 mol%, or 14-15 mol%. In some embodiments, the lipid nanoparticle comprises 1-5% PEG lipid, optionally 1-3 mol%, for example 1.5 to 2.5 mol%, 1-2 mol%, 2-3 mol%, 3-4 mol%, or 4-5 mol%.

In some embodiments, the ionizable cationic lipid is heptadecan-9-yl 8 ((2 hydroxyethyl)(6 oxo 6-(undecyloxy)hexyl)amino)octanoate (Compound 1), the neutral lipid is 1,2 distearoyl-sn-glycero-3-phosphocholine (DSPC), the sterol is cholesterol, and/or the PEG-modified lipid is 1,2 dimyristoyl-sn-glycerol, methoxypolyethyleneglycol (PEG2000 DMG).

In some embodiments, the lipid nanoparticle comprises 20-60 mol% ionizable cationic lipid, 5-25 mol% neutral lipid, 25-55 mol% sterol, and 0.5-15 mol% PEG-modified lipid.

In some embodiments, the lipid nanoparticle comprises: 47 mol% ionizable cationic lipid; 11.5 mol% neutral lipid; 38.5 mol% sterol; and 3.0 mol% PEG-modified lipid; 48 mol% ionizable cationic lipid; 11 mol% neutral lipid; 38.5 mol% sterol; and 2.5 mol% PEG-modified lipid; 49 mol% ionizable cationic lipid; 10.5 mol% neutral lipid; 38.5 mol% sterol; and 2.0 mol% PEG-modified lipid; 50 mol% ionizable cationic lipid; 10 mol% neutral lipid; 38.5 mol% sterol; and 1.5 mol% PEG-modified lipid; or 51 mol% ionizable cationic lipid; 9.5 mol% neutral lipid; 38.5 mol% sterol; and 1.0 mol% PEG-modified lipid.

In some embodiments, the lipid nanoparticle comprises: 47 mol% Compound 1; 11.5 mol% DSPC; 38.5 mol% cholesterol; and 3.0 mol% PEG2000 DMG; 48 mol% Compound 1; 11 mol% DSPC; 38.5 mol% cholesterol; and 2.5 mol% PEG2000 DMG; 49 mol% Compound 1; 10.5 mol% DSPC; 38.5 mol% cholesterol; and 2.0 mol% PEG2000 DMG; 50 mol% Compound 1; 10 mol% DSPC; 38.5 mol% cholesterol; and 1.5 mol% PEG2000 DMG; or 51 mol% Compound 1; 9.5 mol% DSPC; 38.5 mol% cholesterol; and 1.0 mol% PEG2000 DMG.

The entire contents of International Application No. PCT/US2016/058327 (Publication No. WO2017/070626) and International Application No. PCT/US2018/022777 (Publication No. WO2018/170347) are incorporated herein by reference.

SARS-Cov-2

The genome of SARS-CoV-2 is a single-stranded positive-sense RNA (+ssRNA) with the size of 29.8–30 kb encoding about 9860 amino acids (Chan et al.2000, supra; Kim et al. 2020 Cell, May 14; 181(4):914-921.e10.). SARS-CoV-2 is a polycistronic mRNA with 5′-cap and 3′-poly-A tail. The SARS-CoV-2 genome is organized into specific genes encoding structural proteins and nonstructural proteins (Nsps). The order of the structural proteins in the genome is 5′-replicase (open reading frame (ORF)1/ab)-structural proteins [Spike (S)-Envelope (E)-Membrane (M)-Nucleocapsid (N)]-3′. The genome of coronaviruses includes a variable number of open reading frames that encode accessory proteins, nonstructural proteins, and structural proteins (Song et al. 2019 Viruses;11(1):p. 59). Most of the antigenic peptides are located in the structural proteins (Cui et al. 2019 Nat. Rev. Microbiol.;

17(3):181–192). Spike surface glycoprotein (S), a small envelope protein (E), matrix protein (M), and nucleocapsid protein (N) are four main structural proteins. Since S-protein contributes to cell tropism it is capable of inducing neutralizing antibodies (NAb) and protective immunity, it can be considered one of the most important targets in coronavirus vaccine development among all other structural proteins. Moreover, amino acid sequence analysis has shown that S-protein contains conserved regions among the coronaviruses, which may be the basis for universal vaccine development

Antigens

The compositions of the invention, e.g., vaccine compositions, feature nucleic acids, in particular, mRNAs, designed to encode an antigen of interest, e.g., an antigen derived from a betacoronavirus structural protein, in particular, antigens derived from SARS-CoV-2 Spike protein. The compositions of the invention, e.g., vaccine compositions, do not comprise antigens per se, but rather comprise nucleic acids, in particular, mRNA(s) that encode antigens or antigenic sequences once delivered to a cell, tissue or subject. Delivery of nucleic acid molecules, in particular mRNA(s) is achieved by formulating said nucleic acid molecules in appropriate carriers or delivery vehicles (e.g., lipid nanoparticles) such that upon administration to cells, tissues or subjects, nucleic acid is taken up by cells which, in turn, express protein(s) encoded by the nucleic acids, e.g., mRNAs. The term "antigen" as used herein refers to a substance such as a protein (e.g., glycoprotein), polypeptide, peptide, or the like, which elicits an immune response, e.g., elicits an immune response when present in a subject (for example, when present in a human or mammalian subject). The instant invention is based at least in part on the understanding that mRNA-encoded antigens, when expressed from mRNA administered to a cell or subject, can cause the immune system to produce an immune response to the expressed antigen, for example can trigger the production of antibodies against the expresses antigen, e.g., binding and/or neutralizing antibodies, can trigger B and or T cell responses specific to the expressed antigen, and ultimately can cause protective or prophylactic response against subsequent encounter with the antigen or with a pathogen with which the antigen is associated. Preferred mRNA-encoded antigens are “viral antigens”. As used herein, the term “viral antigen” refers to an antigen derived from a virus, for example from a pathogenic virus. The term antigen as used herein can refer to a full-length protein, for example, a full-length viral protein, or can refer to a fragment (e.g., a polypeptide or peptide fragment), subunit or domain of a protein, e.g., a viral protein subunit or domain.

Many proteins have a quaternary or three-dimensional structure, which consists of more than one polypeptide or several polypeptide chains that associate into an oligomeric molecule. As used herein the term "subunit" refers to a single protein molecule, for example, a polypeptide or polypeptide chain resulting from processing of a nascent protein molecule, which subunit assembles (or "coassembles") with other protein molecules (e.g., subunits or chains) to form a protein complex. Proteins can have a relatively small number of subunits and therefore be described as "oligomeric" or can consist of a large number of subunits and therefore be described as "multimeric". The subunits of an oligomeric or multimeric protein may be identical, homologous or totally dissimilar and dedicated to disparate tasks.

Proteins or protein subunits can further comprise domains. As used herein, the term "domain" refers to a distinct functional and/or structural unit within a protein. Typically, a "domain" is responsible for a particular function or interaction, contributing to the overall role of a protein. Domains can exist in a variety of biological contexts. Similar domains (i.e., domains sharing structural, functional and/or sequence homology) can exist within a single protein or can exist within distinct proteins having similar or different functions. A protein domain is often a conserved part of a given protein tertiary structure or sequence that can function and exist independently of the rest of the protein or subunit thereof.

In structural and molecular biology, identical, homologous or similar subunits or domains can help to classify newly identified or novel proteins, as was done immediately upon publication of the SARS-CoV-2 viral genomic sequence.

As used herein, the term antigen is distinct from the term "epitope" which is a substructure of an antigen, e.g., a polypeptide or carbohydrate structure, which may be recognized by an antigen binding site but is insufficient to induce an immune response. The art describes protein antigens that are delivered to subjects or immune cells in isolated form, e.g., isolated protein, polypeptide or peptide antigens, however, the design, testing, validation, and production of protein antigens can be costly and time-consuming, especially when producing proteins at large scale. By contrast, mRNA technology is amenable to rapid design and testing of mRNA constructs encoding a variety of antigens. Moreover, rapid production of mRNA coupled with formulation in appropriate delivery vehicles (e.g., lipid nanoparticles), can proceed quickly and can rapidly produce mRNA vaccines at large scale. Potential benefit also arises from the fact that antigens encoded by the mRNAs of the invention are expressed by the cells of the subject, e.g., are expressed by the human body, and thus the subject, e.g., the human body, serves as the "factory" to produce the antigens which, in turn, elicits the desired immune response.

In preferred aspects, antigens are proteins capable of inducing an immune response (e.g., causing an immune system to produce antibodies against the antigens). Herein, use of the term "antigen" encompasses immunogenic proteins, as well as polypeptides or peptides derived from immunogenic proteins, for example immunogenic fragments (an immunogenic fragment that induces (or is capable of inducing) an immune response to an antigen, unless otherwise stated. It should be understood that the term "protein" encompasses polypeptides and peptides and the term "antigen" encompasses antigenic fragments. Other molecules may be antigenic such as bacterial polysaccharides or combinations of protein and polysaccharide structures, but for the viral vaccines included herein, viral proteins, fragments of viral proteins and designed and or mutated proteins derived from the betacoronavirus SARS-CoV-2 are the antigens featured herein.

Nucleic Acids/mRNA

The vaccine technology described herein features nucleic acids, particularly messenger RNA (mRNA) designed to encode an antigen of interest, e.g., a betacoronavirus spike protein antigen, subunit, domain or fragments (e.g., antigenic fragments) thereof. The nucleic acids, for example mRNAs, of the invention are preferably formulated in appropriate carriers or delivery vehicles (e.g., lipid nanoparticles), such that the nucleic acids, e.g., mRNAs are suitable for use in vivo. When appropriately formulated, nucleic acids, e.g., mRNAs, are capable of being delivered to cells and/or tissues within a subject, e.g., a human subject, to effectuate translation of protein encoded by these nucleic acids.

Nucleic acid molecules are macromolecules comprised of linked nucleotides that carry that carry genetic information and by directing the process of protein synthesis, direct most if not all cellular functions. Nucleic acids comprise a polymer of nucleotides (nucleotide monomers). Thus, nucleic acids are also referred to as polynucleotides (also referred to as polynucleotide chains). The two main classes of nucleic acids are deoxyribonucleic acid (DNA) and ribonucleic acid (RNA). DNA constitutes the genetic material in all free-living organisms and most viruses. RNA is the genetic material of certain viruses, but it is also found in all living cells, where it plays an important role cellular processes, most notably the making of proteins.

Nucleosides are the structural subunit of nucleic acids such as DNA and RNA. A nucleoside is composed of a nitrogenous base (a nucleobase), usually either a pyrimidine (cytosine, thymine or uracil) or a purine (adenine or guanine), covalently attached to a five-carbon carbohydrate ribose or "sugar" which is either ribose or deoxyribose. Nucleotides consist

of a nitrogenous base, a sugar (ribose or deoxyribose) and one to three phosphate groups. In essence, a nucleotide is simply a nucleoside with an additional phosphate group or groups.

The nucleic acid molecules, DNA and RNA, are composed of nucleotides that are linked to one another in a chain by chemical bonds, called ester bonds, between the sugar base of one nucleotide and the phosphate group of the adjacent nucleotide. The sugar is the 3' end, and the phosphate is the 5' end of each nucleotide. The phosphate group attached to the 5' carbon of the sugar on one nucleotide forms an ester bond with the free hydroxyl on the 3' carbon of the next nucleotide. These bonds are called phosphodiester bonds, and the sugar-phosphate backbone is described as extending, or growing, in the 5' to 3' direction when the molecule is synthesized.

The nucleobase portion of nucleic acids features purine bases, adenine (A) and guanine (G), and pyrimidine bases, cytosine (C), thymine (T) in DNA, and uracil (U) in RNA. The sugar portion of nucleic acids features deoxyribose in DNA, ribose in RNA. The five nucleosides are commonly abbreviated to their one-letter codes A, G, C, T and U, respectively. However, thymidine is more commonly written as "dT" ("d" represents "deoxy") as it contains a 2'-deoxyribofuranose moiety rather than the ribofuranose ring found in uridine. This is because thymidine is found in deoxyribonucleic acid (DNA) and not ribonucleic acid (RNA). Conversely, uridine is found in RNA and not DNA. The remaining three nucleosides may be found in both RNA and DNA. In RNA, they would be represented as A, C and G whereas in DNA they would be represented as dA, dC and dG.

The skilled artisan will appreciate that, except where otherwise noted, nucleic acid sequences set forth in the instant application may recite "T"s in a representative DNA sequence but where the sequence represents mRNA, the "T"s would be substituted for "U"s. Thus, any of the DNAs disclosed and identified by a particular sequence identification number herein also disclose the corresponding mRNA sequence complementary to the DNA, where each "T" of the DNA sequence is substituted with "U."

Nucleic acids may be or may include, for example, deoxyribonucleic acids (DNAs), ribonucleic acids (RNAs), e.g. mRNAs, threose nucleic acids (TNAs), glycol nucleic acids (GNAs), peptide nucleic acids (PNAs), locked nucleic acids (LNAs, including LNA having a β-D-ribo configuration, α-LNA having an α-L-ribo configuration (a diastereomer of LNA), 2'-amino-LNA having a 2'-amino functionalization, and 2'-amino- α-LNA having a 2'-amino functionalization), ethylene nucleic acids (ENA), cyclohexenyl nucleic acids (CeNA) and/or chimeras and/or combinations thereof.

Featured in the instant invention are messenger RNAs (mRNAs), particularly mRNAs designed to encode an antigen of interest, e.g., a betacoronavirus spike protein antigen, subunit,

domain or fragments (e.g., antigenic fragments) thereof. Messenger RNA (mRNA), a subtype of RNA, is a single-stranded molecule of RNA that corresponds to the genetic sequence of a gene. mRNA is created during the process of transcription wherein a single strand of DNA is decoded by RNA polymerase, and mRNA is synthesized, i.e., transcribed. mRNA is read by a ribosome in the process of synthesizing a protein, i.e., translation. Accordingly, messenger RNA (mRNA) is an RNA that encodes a (at least one) protein (a naturally-occurring, non-naturally-occurring, or modified polymer of amino acids) and can be translated to produce the encoded protein in vitro, in vivo, in situ, or ex vivo.

The compositions of the present disclosure comprise a (at least one) mRNA having an open reading frame (ORF) encoding a coronavirus antigen. In some embodiments, the mRNA further comprises a 5' UTR, 3' UTR, a poly(A) tail and/or a 5' cap or cap analog. An open reading frame (ORF) is a continuous stretch of DNA or RNA beginning with a start codon (e.g., methionine (ATG or AUG)) and ending with a stop codon (e.g., TAA, TAG or TGA, or UAA, UAG or UGA). An ORF typically encodes a protein. It will be understood that the sequences disclosed herein may further comprise additional elements, e.g., 5′ and 3′ UTRs, but that those elements, unlike the ORF, need not necessarily be present in an mRNA of the present disclosure. It should also be understood that the mRNAs if the invention, e.g., mRNAs featured in the betacoronavirus vaccines of the present disclosure, may include any 5′ untranslated region (UTR) and/or any 3′ UTR. Exemplary UTR sequences are provided in the Sequence Listing (e.g., SEQ ID NOs: 2, 4, 131, and 132); however, other UTR sequences may be used or exchanged for any of the UTR sequences described herein. UTRs may also be omitted from the mRNAs provided herein.

In some embodiments, a composition comprises an mRNA that comprises a nucleotide sequence having at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identity to the nucleotide sequence of any one of SEQ ID NOs: 45, 75, or 90. In some embodiments, a composition comprises an mRNA that comprises a nucleotide sequence having at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identity to the nucleotide sequence of any one of the sequences in Tables 1-15.

In some embodiments, a composition comprises an mRNA that comprises an ORF having at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identity to the nucleotide sequence of any one of SEQ ID NOs: 46, 76, or 91. In some embodiments, a composition comprises an mRNA that comprises an ORF having at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at

least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identity to the nucleotide sequence of any one of the sequences in Table 1-15.

Exemplary sequences of the coronavirus antigens and the mRNA encoding the coronavirus antigens of the compositions of the present disclosure are provided in Tables 1-15.

It should be understood that any one of the antigens encoded by the mRNA described herein may or may not comprise a signal sequence.

Encoded Coronavirus Spike (S) Protein Antigens

The envelope spike (S) proteins of known betacoronaviruses determine the virus host tropism and entry into host cells. Coronavirus spike (S) protein is a choice antigen for the vaccine design as it can induce neutralizing antibodies and protective immunity. S protein is critical for SARS-CoV-2 infection. The organization of the S protein is similar among betacoronaviruses, such as SARS-CoV-2, SARS-CoV, MERS-CoV, HKU1-CoV, MHV-CoV and NL63-CoV.

As used herein, the term "Spike protein" refers to a glycoprotein that that forms homotrimers protruding from the envelope (viral surface) of viruses including betacoronaviruses. Trimerized Spike protein facilitates entry of the virion into a host cell by binding to a receptor on the surface of a host cell followed by fusion of the viral and host cell membranes. The S protein is a highly glycosylated and large type I transmembrane fusion protein that is made up of 1,160 to 1,400 amino acids, depending upon the type of virus. Betacoronavirus Spike proteins comprise between about 1100 to 1500 amino acids and comprise the structure (i.e., the domain composition and organization) as set forth in **FIG. 1**. SARS-CoV-2 spike (S) protein is a choice antigen for the vaccine design as it can induce neutralizing antibodies and protective immunity. mRNAs of the invention are designed to produce SARS-CoV-2 Spike proteins (i.e., encode Spike proteins such that Spike protein is expressed when the mRNA is delivered to a cell or tissue, for example a cell or tissue in a subject), as well as antigenic variants thereof. The skilled artisan will understand that, while an essentially full length or complete Spike protein may be necessary for a virus, e.g., a betacoronavirus, to perform its intended function of facilitating virus entry into a host cell, a certain amount of variation in Spike protein structure and/or sequence is tolerated when seeking primarily to elicit an immune response against Spike protein. For example, minor truncation, e.g., of one to a few, possibly up to 5 or up to 10 amino acids from the N- or C-terminus of the encoded Spike protein, e.g., encoded Spike protein antigen, may be tolerated without changing the antigenic properties of the protein. Likewise, variation (e.g., conservative substitution) of one to a few, possibly up to 5 or up to 10 amino acids (or more) of

the encoded Spike protein, e.g., encoded Spike protein antigen, may be tolerated without changing the antigenic properties of the protein. In exemplary embodiments, a Spike protein, e.g., an encoded Spike protein antigen, has the amino acid sequence set forth in any one of the sequences of Tables 1-15 (e.g., derived from the amino acid sequence set forth as SEQ ID NO: 125). In other embodiments, a Spike protein, e.g., an encoded Spike protein antigen, has no greater than 100, no greater than 90, no greater than 80, no greater than 70, no greater than 60, no greater than 50, no greater than 40, no greater than 30, no greater than 20, no greater than 10, or no greater than 5 amino acid substitutions and/or deletions as compared to (when aligned with) a Spike protein having the amino acid sequence as set forth in any one of the sequences of Tables 1-15 (e.g., derived from the amino acid sequence set forth as SEQ ID NO: 125). Where minor variations are made in encoded Spike protein sequences, the variant preferably has the same activity as the reference Spike protein sequence and/or has the same immune specificity as the reference Spike protein, as determined for example, in immunoassays (e.g., enzyme-linked immunosorbent assays (ELISA assays).

S proteins of coronaviruses can be divided into two important functional subunits, of which include the N-terminal S1 subunit, which forms of the globular head of the S protein, and the C-terminal S2 region that forms the stalk of the protein and is directly embedded into the viral envelope. Upon interaction with a potential host cell, the S1 subunit will recognize and bind to receptors on the host cell, specifically angiotensin-converting enzyme 2 (ACE2) receptors, whereas the S2 subunit, which is the most conserved component of the S protein, will be responsible for fusing the envelope of the virus with the host cell membrane. (See e.g., Shang et al., PLoS Pathog. 2020 Mar; 16(3):e1008392.). Each monomer of trimeric S protein trimer contains the two subunits, S1 and S2, mediating attachment and membrane fusion, respectively. See, e.g., **FIG. 1**. As part of the infection process in vivo, the two subunits are separated from each other by an enzymatic cleavage process. S protein is first cleaved by furin-mediated cleavage at the S1/S2 site in infected cells, In vivo, a subsequent serine protease-mediated cleavage event occurs at the S2′ site within S1. In SARS-CoV2, the S1/S2 cleavage site is at amino acids 676 – TQTNSPRRAR/SVA – 688 (referencing SEQ ID NO: 127). The S2' cleavage site is at amino acids 811 – KPSKR/SFI – 818 (referencing SEQ ID NO: 126).

As used herein, for example in the context of designing SARS-CoV-2 S protein antigens encoded by the nucleic acids, e.g., mRNAs, of the invention, the term "S1 subunit" (e.g., S1 subunit antigen) refers to the N-terminal subunit of the Spike protein beginning at the S protein N-terminus and ending at the S1/S2 cleavage site whereas the term "S2 subunit" (e.g., S2 subunit antigen) refers to the C-terminal subunit of the Spike protein beginning at the S1/S2 cleavage site

and ending at the C-terminus of the Spike protein. As described supra, the skilled artisan will understand that, while an essentially full length or complete Spike protein S1 or S2 subunit may be necessary for receptor binding or membrane fusion, respectively, a certain amount of variation in S1 or S2 structure and/or sequence is tolerated when seeking primarily to elicit an immune response against Spike protein subunits. For example, minor truncation, e.g., of one to a few, possibly up to 4, 5, 6, 7, 8, 9 or 10 amino acids from the N- or C-terminus of the encoded subunit, e.g., encoded S1 or S2 protein antigens, may be tolerated without changing the antigenic properties of the protein. Likewise, variation (e.g., conservative substitution) of one to a few, possibly up to 4, 5, 6, 7, 8, 9 or 10 amino acids (or more) of the encoded Spike protein subunits, e.g., encoded S1 or S2 protein antigen, may be tolerated without changing the antigenic properties of the protein(s). In exemplary embodiments, a Spike protein, e.g., an encoded Spike protein antigen, has the amino acid sequence set forth in any one of the sequences of Tables 1-15 (e.g., derived from the amino acid sequence set forth as SEQ ID NO: 125). In other embodiments, a Spike protein subunit, e.g., an encoded S1 or S2 protein antigen, has no greater than 50, no greater than 40, no greater than 30, no greater than 20, no greater than 10, or no greater than 5 amino acid substitutions and/or deletions as compared to (when aligned with) a Spike protein S1 subunit comprising or consisting of amino acids 1-685 or a Spike protein S2 subunit comprising or consisting of amino acids 686-1273 of the Spike protein having the amino acid sequence as set forth as SEQ ID NO: 125. Where minor variations are made in encoded Spike protein subunit sequences, the variant preferably has the same activity as the reference Spike protein subunit sequence and/or has the same immune specificity as the reference Spike protein subunit, as determined for example, in immunoassays (e.g., enzyme-linked immunosorbent assays (ELISA assays).

The S1 and S2 subunits of the SARS-CoV-2 Spike protein further include domains readily discernable by structure and function, which in turn can be featured in designing antigens to be encoded by the nucleic acid vaccines, in particular, mRNA vaccines of the invention. Within the S1 subunit, domains include the N-terminal domain (NTD) and the receptor-binding domain (RBD), said RBD domain further including a receptor-binding motif (RBM). The wild type S1 subunit also includes a signal peptide (SD), N-terminal to the NTD domain and a first subdomain (SD1) and second subdomain (SD2). Within the S2 subunit, domains include fusion peptide (FP), heptad repeat 1 (HR1), heptad repeat 2 (HR2), transmembrane domain (TM), and cytoplasm domain, also known as cytoplasmic tail (CT) (Lu R. et al., supra; Wan et al., J. Virol. Mar 2020, 94 (7) e00127-20). The HR1 and HR2 domains can be referred to as the "fusion core region" of SARS-CoV-2 (Xia et al., 2020 Cell Mol Immunol. Jan; 17(1):1-12.). **FIG. 1** depicts

the domain architecture in the SARS-CoV-2 Spike protein. The S1 subunit includes an N terminal domain (NTD), a linker region, a receptor binding domain (RBD), a first subdomain (SD1), and a second subdomain (SD2). An S1 subunit may be modified to add a C-terminal transmembrane domain (TM) or it may be soluble. The S2 subunit includes, inter alia, a first heptad repeat (HR1), a second heptad repeat (HR2), a transmembrane domain (TM), and a cytoplasmic tail. A soluble S2 subunit may be generated without a TM domain.

The NTD and RBD of S1 are good antigens for the vaccine design approach of the invention as these domains have been shown to be the targets of neutralizing antibodies in betacoronavirus-infected individuals. As used herein, for example, in the context of an antigen design (said antigen encoded by an mRNA of the invention and to be expressed, for example, from and mRNA vaccine of the invention), the term "N-terminal domain" or "NTD" refers to a domain within the SARS-CoV-2 S1 subunit comprising approximately 290 amino acids in length, having identity to amino acids 1-290 of the S1 subunit of the Spike protein having the amino acid sequence set forth as SEQ ID NO: 125. As used herein, for example, in the context of an antigen design (said antigen encoded by an mRNA of the invention and to be expressed, for example, from and mRNA vaccine of the invention), the term "receptor binding domain" or "RBD" refers to a domain within the S1 subunit of SARS-CoV-2 comprising approximately 175-225 amino acids in length, having identity to amino acids 316-517 of the S1 subunit of the Spike protein having the amino acid sequence set forth as SEQ ID NO: 125. As used herein, the term "receptor binding motif" refers to the portion of the RBD that directly contacts the ACE2 receptor. Expressed RBDs are predicted to specifically bind to angiotensin-converting enzyme 2 (ACE2) as its receptor and/or specifically react with RBD-binding and/or neutralizing antibodies, e.g., CR3022.

The compositions provided herein include mRNA that may encode any one or more full-length or partial (truncated or other deletion of sequence) S protein subunit (e.g., S1 or S2 subunit), one or more domain or combination of domains of an S protein subunit (e.g., NTD, RBD, or NTD-RBD fusions, with or without an SD1 and/or SD2), or chimeras of full-length or partial and S2 protein subunits. Other S protein subunit and/or domain configurations are contemplated herein.

FIG. 2 and FIG. 6 depict exemplary domain and subunit antigens derived from the SARS-CoV-2 Spike protein. FIG. 2A and 2B depict soluble and transmembrane RBD antigens respectively. A transmembrane NTD antigen is shown in FIG. 2C. The domain antigens shown in FIGs. 2D-2F and 2I represent exemplary fusion proteins of NTD and RBD, each with a SP and TM domain. Two of the constructs also have a terminal trafficking domain (CD86 and/or

CD11b). The domains are linked through linkers, in particular GS linkers or a PADRE linker (FIG. 2I). Domain constructs having an RBD domain N-terminal to an NTD domain are depicted in FIGs. 2G and 2H. Each construct may also include a SP and/or TM domain.

Encoded Subunit Antigens

Some aspects of the present disclosure provide compositions comprising an mRNA that encodes a (at least one) subunit of a SARS-CoV-2 S protein. In some embodiments, the mRNA encodes an S1 subunit (e.g., full length or partial). In other embodiments, the mRNA encodes an S2 subunit (e.g., full length or partial). In yet other embodiments, the mRNA encodes a chimeric S1-S2 protein, wherein one of the subunits is from a SARS-CoV-2 S protein, and the other subunit is from another organism, e.g., a virus, such influenza virus. The SARS-CoV-2 subunits (S1 and/or S2) encoded by the mRNA of the present disclosure may be soluble or membrane bound (e.g., linked to a transmembrane domain). Exemplary antigen designs based on S2 are shown in FIG. 6. FIG. 6A depicts a full length S2, including the FP, HR1, HR2, TM and CT domains. A version of S2 comprised of linkers between subunits is shown in FIG. 6B. Domain antigens without the CT domain are shown in FIG. 6C and 6D.

Soluble Subunit Antigens

A soluble protein is present in the cytoplasm of a cell or is secreted from a cell (e.g., not membrane bound). Soluble antigens secreted by cells may be opsonized by complement and captured by follicular dendritic cells in lymph nodes, where they may be recognized by B cells specific to epitopes present on the expressed protein. The expression of subunit antigens further allows focusing of the immune response to specific subunits and with minimal stimulation of memory B and T cells specific to other domains of the antigen that are shared with other related viruses. Without being bound by theory, it is thought that presentation of the SARS-CoV-2 S1 subunit, including the NTD, the RBD, and, in some instances, the intervening polypeptides of the SARS-CoV-2 S1 subunit, in soluble form, generates an S1 subunit-specific immune response. Thus, in some embodiments, an mRNA provided herein encodes a soluble SARS-CoV-2 S1 subunit antigen and/or a soluble SARS-CoV-2 S2 subunit antigen. A non-limiting example of a soluble SARS-CoV-2 S1 subunit antigen and the mRNA encoding it is provided in **Tables 1A** and **1B** below. Other examples of soluble SARS-CoV-2 subunit antigens are provided herein.

Table 1A. Soluble Subunit Antigen

Name	SEQ ID NO:	
	mRNA ORF	Protein
SARS-CoV-2 Soluble S1 Subunit	3	5

Table 1B. Soluble Subunit Antigen

SARS-CoV-2 Soluble S1 Subunit		
SEQ ID NO: 1 consists of from 5' end to 3' end: 5' UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 3 and 3' UTR SEQ ID NO: 4.		1
Chemistry	1-methylpseudouridine	
Cap	7mG(5')ppp(5')NImpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCG GCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGUGAGCAGCCAGUGCG UGAACCUGACCACCCGGACCCAGCUGCCACCAGCCUACACCAACAG CUUCACCCGGGGCGUCUACUACCCCGACAAGGUGUUCCGGAGCAGC GUCCUGCACAGCACCCAGGACCUGUUCCUGCCCUUCUUCAGCAACG UGACCUGGUUCCACGCCAUCCACGUGAGCGGCACCAACGGCACCAA GCGGUUCGACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUC GCCAGCACCGAGAAGAGCAACAUCAUCCGGGGCUGGAUCUUCGGCA CCACCCUGGACAGCAAGACCCAGAGCCUGCUGAUCGUGAAUAACGC CACCAACGUGGUGAUCAAGGUGUGCGAGUUCCAGUUCUGCAACGAC CCCUUCCUGGGCGUGUACUACCACAAGAACAACAAGAGCUGGAUGG AGAGCGAGUUCCGGGUGUACAGCAGCGCCAACAACUGCACCUUCGA GUACGUGAGCCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAGGGC AACUUCAAGAACCUGCGGGAGUUCGUGUUCAAGAACAUCGACGGCU ACUUCAAGAUCUACAGCAAGCACACCCCAAUCAACCUGGUGCGGGA UCUGCCCCAGGGCUUCUCAGCCCUGGAGCCCCUGGUGGACCUGCCC AUCGGCAUCAACAUCACCCGGUUCCAGACCCUGCUGGCCCUGCACC GGAGCUACCUGACCCCAGGCGACAGCAGCAGCGGGUGGACAGCAGG CGCGGCUGCUUACUACGUGGGCUACCUGCAGCCCCGGACCUUCCUG CUGAAGUACAACGAGAACGGCACCAUCACCGACGCCGUGGACUGCG CCCUGGACCCUCUGAGCGAGACCAAGUGCACCCUGAAGAGCUUCAC CGUGGAGAAGGGCAUCUACCAGACCAGCAACUUCCGGGUGCAGCCC ACCGAGAGCAUCGUGCGGUUCCCCAACAUCACCAACCUGUGCCCCU UCGGCGAGGUGUUCAACGCCACCCGGUUCGCCAGCGUGUACGCCUG GAACCGGAAGCGGAUCAGCAACUGCGUGGCCGACUACAGCGUGCUG UACAACAGCGCCAGCUUCAGCACCUUCAAGUGCUACGGCGUGAGCC CCACCAAGCUGAACGACCUGUGCUUCACCAACGUGUACGCCGACAG CUUCGUGAUCCGUGGCGACGAGGUGCGGCAGAUCGCACCCGGCCAG ACAGGCAAGAUCGCCGACUACAACUACAAGCUGCCCGACGACUUCA CCGGCUGCGUGAUCGCCUGGAACAGCAACAACCUCGACAGCAAGGU GGGCGGCAACUACAACUACCUGUACCGGCUGUUCCGGAAGAGCAAC CUGAAGCCCUUCGAGCGGGACAUCAGCACCGAGAUCUACCAAGCCG GCUCCACCCCUUGCAACGGCGUGGAGGGCUUCAACUGCUACUUCCC UCUGCAGAGCUACGGCUUCCAGCCCACCAACGGCGUGGGCUACCAG CCCUACCGGGUGGUGGUGCUGAGCUUCGAGCUGCUGCACGCCCCAG CCACCGUGUGUGGCCCCAAGAAGAGCACCAACCUGGUGAAGAACAA GUGCGUGAACUUCAACUUCAACGGCCUUACCGGCACCGGCGUGCUG ACCGAGAGCAACAAGAAAUUCCUGCCCUUUCAGCAGUUCGGCCGGG ACAUCGCCGACACCACCGACGCUGUGCGGGAUCCCCAGACCCUGGA GAUCCUGGACAUCACCCCUUGCAGCUUCGGCGGCGUGAGCGUGAUC ACCCCAGGCACCAACACCAGCAACCAGGUGGCCGUGCUGUACCAGG ACGUGAACUGCACCGAGGUGCCCGUGGCCAUCCACGCCGACCAGCU GACACCCACCUGGCGGGUCUACAGCACCGGCAGCAACGUGUUCCAG ACCCGGGCCGGUUGCCUGAUCGGCGCCGAGCACGUGAACAACAGCU	3

	ACGAGUGCGACAUCCCCAUCGGCGCCGGCAUCUGUGCCAGCUACCA GACCCAGACCAAUUCA	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGG CCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGC	4
Corresponding amino acid sequence	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSS VLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYF ASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCND PFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQG NFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLP IGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFL LKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQP TESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVL YNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQ TGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSN LKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQ PYRVVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVL TESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVI TPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQ TRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNS	5
PolyA tail	100 nt	

Membrane Bound Subunit Antigens

A membrane bound protein is anchored in a cell membrane (not soluble). Without being bound by theory, it is thought that antigen presenting cells will carry the embedded antigen to the draining lymph nodes to generate a strong immune response. The germinal center reaction that occurs in the draining lymph node involves prolonged contact between $CD4^+$ T_{FH} cells and B cells, allowing co-stimulation and local cytokine signals such as IL-4 and IL-21 that favor replication of B cells specific to the presented antigen and class switching to the production of IgG1, each of which may promote the generation of long-lived plasma cells and memory B cells. Thus, in some embodiments, an mRNA encodes a membrane bound SARS-CoV-2 S1 subunit antigen and/or a membrane bound SARS-CoV-2 S2 subunit antigen. In some embodiments, a membrane bound antigen (e.g., S1 subunit, S2 subunit, NTD, RBD, or any combination thereof) is linked to a transmembrane domain, e.g., a naturally occurring transmembrane domain or a heterologous transmembrane domain (derived from a heterologous protein), which is responsible for anchoring the protein in the cell membrane. A non-limiting example of a membrane bound SARS-CoV-2 S1 subunit antigen and a SARS-CoV-2 S2 subunit antigen and the mRNA encoding them are provided in **Tables 2A** and **2B** below. Other membrane bound SARS-CoV-2 S1 subunit antigens are contemplated herein.

Table 2A. Membrane Bound Subunit Antigen

Name	SEQ ID NO:	
	mRNA ORF	Protein
SARS-CoV-2 S1 Subunit Linked to Transmembrane Domain (S1-666-TM)	16	17
SARS-CoV-2 S2 Subunit Linked to Transmembrane Domain (S2-TM)	145	146

Table 2B. Membrane Bound Subunit Antigen

SARS-CoV-2 S1 Subunit Linked to Transmembrane Domain (S1-666-TM)		
SEQ ID NO: 15 consists of from 5' end to 3' end: 5' UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 16 and 3' UTR SEQ ID NO: 4.		15
Chemistry	1-methylpseudouridine	
Cap	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCG GCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGUGAGCAGCCAGUGCG UGAACCUGACCACCCGGACCCAGCUGCCACCAGCCUACACCAACAG CUUCACCCGGGGCGUCUACUACCCCGACAAGGUGUUCCGGAGCAGC GUCCUGCACAGCACCCAGGACCUGUUCCUGCCCUUCUUCAGCAACG UGACCUGGUUCCACGCCAUCCACGUGAGCGGCACCAACGGCACCAA GCGGUUCGACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUC GCCAGCACCGAGAAGAGCAACAUCAUCCGGGGCUGGAUCUUCGGCA CCACCCUGGACAGCAAGACCCAGAGCCUGCUGAUCGUGAAUAACGC CACCAACGUGGUGAUCAAGGUGUGCGAGUUCCAGUUCUGCAACGAC CCCUUCCUGGGCGUGUACUACCACAAGAACAACAAGAGCUGGAUGG AGAGCGAGUUCCGGGUGUACAGCAGCGCCAACAACUGCACCUUCGA GUACGUGAGCCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAGGGC AACUUCAAGAACCUGCGGGAGUUCGUGUUCAAGAACAUCGACGGCU ACUUCAAGAUCUACAGCAAGCACACCCCAAUCAACCUGGUGCGGGA UCUGCCCCAGGGCUUCUCAGCCCUGGAGCCCCUGGUGGACCUGCCC AUCGGCAUCAACAUCACCCGGUUCCAGACCCUGCUGGCCCUGCACC GGAGCUACCUGACCCCAGGCGACAGCAGCAGCGGGUGGACAGCAGG CGCGGCUGCUUACUACGUGGGCUACCUGCAGCCCCGGACCUUCCUG CUGAAGUACAACGAGAACGGCACCAUCACCGACGCCGUGGACUGCG CCCUGGACCCUCUGAGCGAGACCAAGUGCACCCUGAAGAGCUUCAC CGUGGAGAAGGGCAUCUACCAGACCAGCAACUUCCGGGUGCAGCCC ACCGAGAGCAUCGUGCGGUUCCCCAACAUCACCAACCUGUGCCCCU UCGGCGAGGUGUUCAACGCCACCCGGUUCGCCAGCGUGUACGCCUG GAACCGGAAGCGGAUCAGCAACUGCGUGGCCGACUACAGCGUGCUG UACAACAGCGCCAGCUUCAGCACCUUCAAGUGCUACGGCGUGAGCC CCACCAAGCUGAACGACCUGUGCUUCACCAACGUGUACGCCGACAG CUUCGUGAUCCGUGGCGACGAGGUGCGGCAGAUCGCACCCGGCCAG ACAGGCAAGAUCGCCGACUACAACUACAAGCUGCCCGACGACUUCA CCGGCUGCGUGAUCGCCUGGAACAGCAACAACCUCGACAGCAAGGU GGGCGGCAACUACAACUACCUGUACCGGCUGUUCCGGAAGAGCAAC CUGAAGCCCUUCGAGCGGGACAUCAGCACCGAGAUCUACCAAGCCG GCUCCACCCCUUGCAACGGCGUGGAGGGCUUCAACUGCUACUUCCC UCUGCAGAGCUACGGCUUCCAGCCCACCAACGGCGUGGGCUACCAG CCCUACCGGGUGGUGGUGCUGAGCUUCGAGCUGCUGCACGCCCCAG CCACCGUGUGUGGCCCCAAGAAGAGCACCAACCUGGUGAAGAACAA GUGCGUGAACUUCAACUUCAACGGCCUUACCGGCACCGGCGUGCUG ACCGAGAGCAACAAGAAAUUCCUGCCCUUUCAGCAGUUCGGCCGGG ACAUCGCCGACACCACCGACGCUGUGCGGGAUCCCCAGACCCUGGA GAUCCUGGACAUCACCCCUUGCAGCUUCGGCGGCGUGAGCGUGAUC	16

	ACCCCAGGCACCAACACCAGCAACCAGGUGGCCGUGCUGUACCAGG ACGUGAACUGCACCGAGGUGCCCGUGGCCAUCCACGCCGACCAGCU GACACCCACCUGGCGGGUCUACAGCACCGGCAGCAACGUGUUCCAG ACCCGGGCCGGUUGCCUGAUCGGCGCCGAGCACGUGAACAACAGCU ACGAGUGCGACAUCCCCAUCGGCGCCGGCAUCUGUGCCAGCUACCA GACCCAGACCAAUUCAUCUGGCGGAGGCAGCAUCCUGGCCAUCUAC AGCACCGUGGCCAGCAGCCUGGUGCUGCUGGUGAGCCUGGGCGCCA UCAGCUUC	
3' UTR	UGAUAAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGG CCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGC	4
Corresponding amino acid sequence	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSS VLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYF ASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCND PFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQG NFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLP IGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFL LKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQP TESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVL YNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQ TGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSN LKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQ PYRVVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVL TESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVI TPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQ TRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSGGGSILAIYS TVASSLVLLVSLGAISF	17
PolyA tail	100 nt	
SARS-CoV-2 S2 Subunit comprising a Transmembrane Domain (S2-TM)		
SEQ ID NO: 147 consists of from 5' end to 3' end: 5' UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 145 and 3' UTR SEQ ID NO: 4.		147
Chemistry	1-methylpseudouridine	
Cap	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCG GCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUACAGCAUGCAGCUGGCUAGCUGCGUGACCCUGACCCUGGUGC UGCUGGUGAACAGCCAGGGCGCCGAGAACAGCGUGGCCUACAGCAA CAACAGCAUCGCCAUCCCCACCAACUUCACCAUCAGCGUGACCACC GAGAUUCUGCCCGUGAGCAUGACCAAGACCAGCGUGGACUGCACCA UGUACAUCUGCGGCGACAGCACCGAGUGCAGCAACCUGCUGCUGCA GUACGGCAGCUUCUGCACCCAGCUGAACCGGGCCCUGACCGGCAUC GCCGUGGAGCAGGACAAGAACACCCAGGAGGUGUUCGCCCAGGUGA AGCAGAUCUACAAGACCCCUCCCAUCAAGGACUUCGGCGGCUUCAA CUUCAGCCAGAUCCUGCCCGACCCCAGCAAGCCCAGCAAGCGGAGC UUCAUCGAGGACCUGCUGUUCAACAAGGUGACCCUAGCCGACGCCG GCUUCAUCAAGCAGUACGGCGACUGCCUCGGCGACAUAGCCGCCCG GGACCUGAUCUGCGCCCAGAAGUUCAACGGCCUGACCGUGCUGCCU CCCCUGCUGACCGACGAGAUGAUCGCCCAGUACACCAGCGCCCUGU UAGCCGGAACCAUCACCAGCGGCUGGACUUUCGGCGCUGGAGCCGC UCUGCAGAUCCCCUUCGCCAUGCAGAUGGCCUACCGGUUCAACGGC AUCGGCGUGACCCAGAACGUGCUGUACGAGAACCAGAAGCUGAUCG CCAACCAGUUCAACAGCGCCAUCGGCAAGAUCCAGGACAGCCUGAG CAGCACCGCUAGCGCCCUGGGCAAGCUGCAGGACGUGGUGAACCAG AACGCCCAGGCCCUGAACACCCUGGUGAAGCAGCUGAGCAGCAACU UCGGCGCCAUCAGCAGCGUGCUGAACGACAUCCUGAGCCGGCUGGA CCCUCCCGAGGCCGAGGUGCAGAUCGACCGGCUGAUCACUGGCCGG CUGCAGAGCCUGCAGACCUACGUGACCCAGCAGCUGAUCCGGGCCG CCGAGAUUCGGGCCAGCGCCAACCUGGCCGCCACCAAGAUGAGCGA GUGCGUGCUGGGCCAGAGCAAGCGGGUGGACUUCUGCGGCAAGGGC UACCACCUGAUGAGCUUUCCCCAGAGCGCACCCCACGGAGUGGUGU	145

	UCCUGCACGUGACCUACGUGCCCGCCCAGGAGAAGAACUUCACCAC CGCCCCAGCCAUCUGCCACGACGGCAAGGCCCACUUUCCCCGGGAG GGCGUGUUCGUGAGCAACGGCACCCACUGGUUCGUGACCCAGCGGA ACUUCUACGAGCCCCAGAUCAUCACCACCGACAACACCUUCGUGAG CGGCAACUGCGACGUGGUGAUCGGCAUCGUGAACAACACCGUGUAC GAUCCCCUGCAGCCCGAGCUGGACAGCUUCAAGGAGGAGCUGGACA AGUACUUCAAGAAUCACACCAGCCCCGACGUGGACCUGGGCGACAU CAGCGGCAUCAACGCCAGCGUGGUGAACAUCCAGAAGGAGAUCGAU CGGCUGAACGAGGUGGCCAAGAACCUGAACGAGAGCCUGAUCGACC UGCAGGAGCUGGGCAAGUACGAGCAGUACAUCAAGUGGCCCUGGUA CAUCUGGCUGGGCUUCAUCGCCGGCCUGAUCGCCAUCGUGAUGGUG ACCAUCAUGCUGUGCUGCAUGACCAGCUGCUGCAGCUGCCUGAAGG GCUGUUGCAGCUGCGGCAGCUGCUGCAAGUUCGACGAGGACGACAG CGAGCCCGUGCUGAAGGGCGUGAAGCUGCACUACACC	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGG CCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGC	4
Corresponding amino acid sequence	MYSMQLASCVTLTLVLLVNSQGAENSVAYSNNSIAIPTNFTISVTT EILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGI AVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRS FIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLP PLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNG IGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNQ NAQALNTLVKQLSSNFGAISSVLNDILSRLDPPEAEVQIDRLITGR LQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKG YHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPRE GVFVSNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVY DPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEID RLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGLIAIVMV TIMLCCMTSCCSCLKGCCSCGSCCKFDEDDSEPVLKGVKLHYT	146
PolyA tail	100 nt	

Subunit Antigen Truncations and RBD Deletions

In some embodiments, a composition comprises an mRNA that encodes an S1 subunit that has been modified to remove the RBD or a portion of the RBD. Truncation of the S1 subunit provides fewer epitopes for the immune system to recognize, thereby biasing the immune response to the remaining epitopes, which may select for antibodies to specific epitopes that are important for virus neutralization. Truncation or partial deletion of the RBD may prevent the expressed protein or cells carrying it from interacting with receptor ACE2, making it more likely to reach the lymph node and stimulate a desired immune response. Furthermore, removing the RBD may prevent epitope masking by cross-reactive antibodies previously raised against related viruses, and thus focus the elicited immune response toward the desired antigen specifically. Additionally, removal of the RBD may alter the conformation of the expressed subunit, allowing B cells specific to these alternative conformational epitopes to uptake and present linear peptides to T cells, thereby indirectly enhancing the $CD4^+$ T cell response to those epitopes, which are still present in the native conformation.

In some embodiments, a composition comprises an mRNA that encodes an S1 subunit that has been modified to remove the RBD or a portion of the RBD, wherein the S2 subunit

contains a glycan. Glycans are attached to proteins by N-linked glycosylation via asparagine residues or O-linked glycosylation on serine or threonine residues. The presence of a glycan shield on some components of a protein may mask peptide epitopes, thereby focusing the antibody response towards other exposed peptide epitopes. Furthermore, glycosylated proteins also elicit antibodies that recognize the coating glycans. B cells that recognize the glycan epitope will intake and present linear peptide epitopes to $CD4^+$ T cells, thereby boosting the $CD4^+$ T cell response to linear epitopes found throughout the protein.

Non-limiting examples of truncated SARS-CoV-2 S1 subunit antigens and the mRNA encoding them are provided in **Tables 3A** and **3B** below.

Non-limiting examples of SARS-CoV-2 S1 subunits having an RBD deletion and the mRNA encoding them are provided in **Tables 4A** and **4B** below.

Table 3A. Subunit Antigen Truncations

Name	SEQ ID NO:	
	mRNA ORF	Protein
SARS-CoV-2 S1 Subunit Truncated and Linked to Transmembrane Domain (S1-531-TM)	19	20
SARS-CoV-2 S1 Subunit Truncated and Linked to Transmembrane Domain (S1-594-TM)	22	23
SARS-CoV-2 S1 Subunit Truncated with PolyG and Linked to Transmembrane Domain (S1-594-PolyG-TM)	25	26
SARS-CoV-2 S1 Subunit Truncated with PolyG/DS and Linked to Transmembrane Domain (S1-594-PolyG-DS-TM)	28	29
SARS-CoV-2 S1 Subunit Truncated and Linked to Transmembrane Domain (S1-666-TM)	150	151

Table 3B. Subunit Antigen Truncations

SARS-CoV-2 S1 Subunit Truncated and Linked to Transmembrane Domain (S1-531-TM)		
SEQ ID NO: 18 consists of from 5' end to 3' end: 5' UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 19 and 3' UTR SEQ ID NO: 4.		18
Chemistry	1-methylpseudouridine	
Cap	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCG GCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGUGAGCAGCCAGUGCG UGAACCUGACCACCCGGACCCAGCUGCCACCAGCCUACACCAACAG CUUCACCCGGGGCGUCUACUACCCCGACAAGGUGUUCCGGAGCAGC GUCCUGCACAGCACCCAGGACCUGUUCCUGCCCUUCUUCAGCAACG UGACCUGGUUCCACGCCAUCCACGUGAGCGGCACCAACGGCACCAA GCGGUUCGACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUC GCCAGCACCGAGAAGAGCAACAUCAUCCGGGGCUGGAUCUUCGGCA CCACCCUGGACAGCAAGACCCAGAGCCUGCUGAUCGUGAAUAACGC	19

	CACCAACGUGGUGAUCAAGGUGUGCGAGUUCCAGUUCUGCAACGAC CCCUUCCUGGGCGUGUACUACCACAAGAACAACAAGAGCUGGAUGG AGAGCGAGUUCCGGGUGUACAGCAGCGCCAACAACUGCACCUUCGA GUACGUGAGCCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAGGGC AACUUCAAGAACCUGCGGGAGUUCGUGUUCAAGAACAUCGACGGCU ACUUCAAGAUCUACAGCAAGCACACCCCAAUCAACCUGGUGCGGGA UCUGCCCCAGGGCUUCUCAGCCCUGGAGCCCCUGGUGGACCUGCCC AUCGGCAUCAACAUCACCCGGUUCCAGACCCUGCUGGCCCUGCACC GGAGCUACCUGACCCCAGGCGACAGCAGCAGCGGGUGGACAGCAGG CGCGGCUGCUUACUACGUGGGCUACCUGCAGCCCCGGACCUUCCUG CUGAAGUACAACGAGAACGGCACCAUCACCGACGCCGUGGACUGCG CCCUGGACCCUCUGAGCGAGACCAAGUGCACCCUGAAGAGCUUCAC CGUGGAGAAGGGCAUCUACCAGACCAGCAACUUCCGGGUGCAGCCC ACCGAGAGCAUCGUGCGGUUCCCCAACAUCACCAACCUGUGCCCCU UCGGCGAGGUGUUCAACGCCACCCGGUUCGCCAGCGUGUACGCCUG GAACCGGAAGCGGAUCAGCAACUGCGUGGCCGACUACAGCGUGCUG UACAACAGCGCCAGCUUCAGCACCUUCAAGUGCUACGGCGUGAGCC CCACCAAGCUGAACGACCUGUGCUUCACCAACGUGUACGCCGACAG CUUCGUGAUCCGUGGCGACGAGGUGCGGCAGAUCGCACCCGGCCAG ACAGGCAAGAUCGCCGACUACAACUACAAGCUGCCCGACGACUUCA CCGGCUGCGUGAUCGCCUGGAACAGCAACAACCUCGACAGCAAGGU GGGCGGCAACUACAACUACCUGUACCGGCUGUUCCGGAAGAGCAAC CUGAAGCCCUUCGAGCGGGACAUCAGCACCGAGAUCUACCAAGCCG GCUCCACCCCUUGCAACGGCGUGGAGGGCUUCAACUGCUACUUCCC UCUGCAGAGCUACGGCUUCCAGCCCACCAACGGCGUGGGCUACCAG CCCUACCGGGUGGUGGUGCUGAGCUUCGAGCUGCUGCACGCCCCAG CCACCGUGUGUGGCCCCAAGAAGAGCACCUCUGGCGGAGGCAGCAU CCUGGCCAUCUACAGCACCGUGGCCAGCAGCCUGGUGCUGCUGGUG AGCCUGGGCGCCAUCAGCUUC	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGG CCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGC	4
Corresponding amino acid sequence	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSS VLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYF ASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCND PFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQG NFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLP IGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFL LKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQP TESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVL YNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQ TGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSN LKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQ PYRVVVLSFELLHAPATVCGPKKSTsgggsilaiystvasslvllv slgaisf	20
PolyA tail	100 nt	
SARS-CoV-2 S1 Subunit Truncated and Linked to Transmembrane Domain (S1-594-TM)		
SEQ ID NO: 21 consists of from 5' end to 3' end: 5' UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 22 and 3' UTR SEQ ID NO: 4.		21
Chemistry	1-methylpseudouridine	
Cap	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCG GCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGUGAGCAGCCAGUGCG UGAACCUGACCACCCGGACCCAGCUGCCACCAGCCUACACCAACAG CUUCACCCGGGGCGUCUACUACCCCGACAAGGUGUUCCGGAGCAGC GUCCUGCACAGCACCCAGGACCUGUUCCUGCCCUUCUUCAGCAACG UGACCUGGUUCCACGCCAUCCACGUGAGCGGCACCAACGGCACCAA GCGGUUCGACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUC GCCAGCACCGAGAAGAGCAACAUCAUCCGGGGCUGGAUCUUCGGCA	22

	CCACCCUGGACAGCAAGACCCAGAGCCUGCUGAUCGUGAAUAACGC CACCAACGUGGUGAUCAAGGUGUGCGAGUUCCAGUUCUGCAACGAC CCCUUCCUGGGCGUGUACUACCACAAGAACAACAAGAGCUGGAUGG AGAGCGAGUUCCGGGUGUACAGCAGCGCCAACAACUGCACCUUCGA GUACGUGAGCCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAGGGC AACUUCAAGAACCUGCGGGAGUUCGUGUUCAAGAACAUCGACGGCU ACUUCAAGAUCUACAGCAAGCACACCCCAAUCAACCUGGUGCGGGA UCUGCCCCAGGGCUUCUCAGCCCUGGAGCCCCUGGUGGACCUGCCC AUCGGCAUCAACAUCACCCGGUUCCAGACCCUGCUGGCCCUGCACC GGAGCUACCUGACCCCAGGCGACAGCAGCAGCGGGUGGACAGCAGG CGCGGCUGCUUACUACGUGGGCUACCUGCAGCCCCGGACCUUCCUG CUGAAGUACAACGAGAACGGCACCAUCACCGACGCCGUGGACUGCG CCCUGGACCCUCUGAGCGAGACCAAGUGCACCCUGAAGAGCUUCAC CGUGGAGAAGGGCAUCUACCAGACCAGCAACUUCCGGGUGCAGCCC ACCGAGAGCAUCGUGCGGUUCCCCAACAUCACCAACCUGUGCCCCU UCGGCGAGGUGUUCAACGCCACCCGGUUCGCCAGCGUGUACGCCUG GAACCGGAAGCGGAUCAGCAACUGCGUGGCCGACUACAGCGUGCUG UACAACAGCGCCAGCUUCAGCACCUUCAAGUGCUACGGCGUGAGCC CCACCAAGCUGAACGACCUGUGCUUCACCAACGUGUACGCCGACAG CUUCGUGAUCCGUGGCGACGAGGUGCGGCAGAUCGCACCCGGCCAG ACAGGCAAGAUCGCCGACUACAACUACAAGCUGCCCGACGACUUCA CCGGCUGCGUGAUCGCCUGGAACAGCAACAACCUCGACAGCAAGGU GGGCGGCAACUACAACUACCUGUACCGGCUGUUCCGGAAGAGCAAC CUGAAGCCCUUCGAGCGGGACAUCAGCACCGAGAUCUACCAAGCCG GCUCCACCCCUUGCAACGGCGUGGAGGGCUUCAACUGCUACUUCCC UCUGCAGAGCUACGGCUUCCAGCCCACCAACGGCGUGGGCUACCAG CCCUACCGGGUGGUGGUGCUGAGCUUCGAGCUGCUGCACGCCCCAG CCACCGUGUGUGGCCCCAAGAAGAGCACCAACCUGGUGAAGAACAA GUGCGUGAACUUCAACUUCAACGGCCUUACCGGCACCGGCGUGCUG ACCGAGAGCAACAAGAAAUUCCUGCCCUUUCAGCAGUUCGGCCGGG ACAUCGCCGACACCACCGACGCUGUGCGGGAUCCCCAGACCCUGGA GAUCCUGGACAUCACCCCUUGCAGCUUCGGCGGCUCUGGCGGAGGC AGCAUCCUGGCCAUCUACAGCACCGUGGCCAGCAGCCUGGUGCUGC UGGUGAGCCUGGGCGCCAUCAGCUUC	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGG CCUCCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGC	4
Corresponding amino acid sequence	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSS VLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYF ASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCND PFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQG NFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLP IGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFL LKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQP TESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVL YNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQ TGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSN LKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQ PYRVVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVL TESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGsggg silaiystvasslvllvslgaisf	23
PolyA tail	100 nt	
SARS-CoV-2 S1 Subunit Truncated with PolyG and Linked to Transmembrane Domain (S1-594-PolyG-TM)		
SEQ ID NO: 24 consists of from 5' end to 3' end: 5' UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 25 and 3' UTR SEQ ID NO: 4.		24
Chemistry	1-methylpseudouridine	
Cap	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCG GCGCCGCCACC	2

ORF of mRNA Construct (excluding the stop codon)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGUGAGCAGCCAGUGCG UGAACCUGACCACCCGGACCCAGCUGCCACCAGCCUACACCAACAG CUUCACCCGGGGCGUCUACUACCCCGACAAGGUGUUCCGGAGCAGC GUCCUGCACAGCACCCAGGACCUGUUCCUGCCCUUCUUCAGCAACG UGACCUGGUUCCACGCCAUCCACGUGAGCGGCACCAACGGCACCAA GCGGUUCGACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUC GCCAGCACCGAGAAGAGCAACAUCAUCCGGGGCUGGAUCUUCGGCA CCACCCUGGACAGCAAGACCCAGAGCCUGCUGAUCGUGAAUAACGC CACCAACGUGGUGAUCAAGGUGUGCGAGUUCCAGUUCUGCAACGAC CCCUUCCUGGGCGUGUACUACCACAAGAACAACAAGAGCUGGAUGG AGAGCGAGUUCCGGGUGUACAGCAGCGCCAACAACUGCACCUUCGA GUACGUGAGCCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAGGGC AACUUCAAGAACCUGCGGGAGUUCGUGUUCAAGAACAUCGACGGCU ACUUCAAGAUCUACAGCAAGCACACCCCAAUCAACCUGGUGCGGGA UCUGCCCCAGGGCUUCUCAGCCCUGGAGCCCCUGGUGGACCUGCCC AUCGGCAUCAACAUCACCCGGUUCCAGACCCUGCUGGCCCUGCACC GGAGCUACCUGACCCCAGGCGACAGCAGCAGCGGGUGGACAGCAGG CGCGGCUGCUUACUACGUGGGCUACCUGCAGCCCCGGACCUUCCUG CUGAAGUACAACGAGAACGGCACCAUCACCGACGCCGUGGACUGCG CCCUGGACCCUCUGAGCGAGACCAAGUGCACCCUGAAGAGCUUCAC CGUGGGCAGCGGCGGCGGCAGCGGCGGAGGCAGCGGAGGAGGCAGC GGCGGAGGCAGUGGAGGCCAGCCCACCGAGAGCAUCGUGCGGUUCC CCAACAUCACCAACCUGUGCCCCUUCGGCGAGGUGUUCAACGCCAC CCGGUUCGCCAGCGUGUACGCCUGGAACCGGAAGCGGAUCAGCAAC UGCGUGGCCGACUACAGCGUGCUGUACAACAGCGCCAGCUUCAGCA CCUUCAAGUGCUACGGCGUGAGCCCCACCAAGCUGAACGACCUGUG CUUCACCAACGUGUACGCCGACAGCUUCGUGAUCCGUGGCGACGAG GUGCGGCAGAUCGCACCCGGCCAGACAGGCAAGAUCGCCGACUACA ACUACAAGCUGCCCCGACGACUUCACCGGCUGCGUGAUCGCCUGGAA CAGCAACAACCUCGACAGCAAGGUGGGCGGCAACUACAACUACCUG UACCGGCUGUUCCGGAAGAGCAACCUGAAGCCCUUCGAGCGGGACA UCAGCACCGAGAUCUACCAAGCCGGCUCCACCCCUUGCAACGGCGU GGAGGGCUUCAACUGCUACUUCCCUCUGCAGAGCUACGGCUUCCAG CCCACCAACGGCGUGGGCUACCAGCCCUACCGGGUGGUGGUGCUGA GCUUCGAGCUGCUGCACGCCCCAGCCACCGUGUGUGGCCCCAAGAA GAGCACCAACCUGGUGAAGAACAAGUGCGUGAACUUCAACUUCAAC GGCCUUACCGGCACCGGCGUGCUGACCGAGAGCAACAAGAAAUUCC UGCCCUUUCAGCAGUUCGGCCGGGACAUCGCCGACACCACCGACGC UGUGCGGGAUCCCCAGACCCUGGAGAUCCUGGACAUCACCCCUUGC AGCUUCGGCGGCUCUGGCGGAGGCAGCAUCCUGGCCAUCUACAGCA CCGUGGCCAGCAGCCUGGUGCUGCUGGUGAGCCUGGGCGCCAUCAG CUUC	25
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGG CCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGC	4
Corresponding amino acid sequence	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSS VLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYF ASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCND PFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQG NFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLP IGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFL LKYNENGTITDAVDCALDPLSETKCTLKSFTVGSGGGSGGGSGGGS GGGSGGQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISN CVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDE VRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYL YRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQ PTNGVGYQPYRVVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFN GLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPC SFGGsgggsilaiystvasslvllvslgaisf	26
PolyA tail	100 nt	
SARS-CoV-2 S1 Subunit Truncated with PolyG/DS and Linked to Transmembrane Domain (S1-594-PolyG-DS-TM)		

SEQ ID NO: 27 consists of from 5' end to 3' end: 5' UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 28 and 3' UTR SEQ ID NO: 4.		27
Chemistry	1-methylpseudouridine	
Cap	7mG(5')ppp(5')NImpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCG GCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGUGAGCAGCCAGUGCG UGAACCUGACCACCCGGACCCAGCUGCCACCAGCCUACACCAACAG CUUCACCCGGGGCGUCUACUACCCCGACAAGGUGUGCCGGAGCAGC GUCCUGCACAGCACCCAGGACCUGUUCCUGCCCUUCUUCAGCAACG UGACCUGGUUCCACGCCAUCCACGUGAGCGGCACCAACGGCACCAA GCGGUUCGACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUC GCCAGCACCGAGAAGAGCAACAUCAUCCGGGGCUGGAUCUUCGGCA CCACCCUGGACAGCAAGACCCAGAGCCUGCUGAUCGUGAAUAACGC CACCAACGUGGUGAUCAAGGUGUGCGAGUUCCAGUUCUGCAACGAC CCCUUCCUGGGCGUGUACUACCACAAGAACAACAAGAGCUGGAUGG AGAGCGAGUUCCGGGUGUACAGCAGCGCCAACAACUGCACCUUCGA GUACGUGAGCCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAGGGC AACUUCAAGAACCUGCGGGAGUUCGUGUUCAAGAACAUCGACGGCU ACUUCAAGAUCUACAGCAAGCACACCCCAAUCAACCUGGUGCGGGA UCUGCCCCAGGGCUUCUCAGCCCUGGAGCCCCUGGUGGACCUGCCC AUCGGCAUCAACAUCACCCGGUUCCAGACCCUGCUGGCCCUGCACC GGAGCUACCUGACCCCAGGCGACAGCAGCAGCGGGUGGACAGCAGG CGCGGCUGCUUACUACGUGGGCUACCUGCAGCCCCGGACCUUCCUG CUGAAGUACAACGAGAACGGCACCAUCACCGACGCCGUGGACUGCG CCCUGGACCCUCUGAGCGAGACCAAGUGCACCCUGAAGAGCUUCAC CGUGGGCAGCGGCGGCGGCAGCGGCGGAGGCAGCGGAGGAGGCAGC GGCGGAGGCAGUGGAGGCCAGCCCACCGAGAGCAUCGUGCGGUUCC CCAACAUCACCAACCUGUGCCCCUUCGGCGAGGUGUUCAACGCCAC CCGGUUCGCCAGCGUGUACGCCUGGAACCGGAAGCGGAUCAGCAAC UGCGUGGCCGACUACAGCGUGCUGUACAACAGCGCCAGCUUCAGCA CCUUCAAGUGCUACGGCGUGAGCCCCACCAAGCUGAACGACCUGUG CUUCACCAACGUGUACGCCGACAGCUUCGUGAUCCGUGGCGACGAG GUGCGGCAGAUCGCACCCGGCCAGACAGGCAAGAUCGCCGACUACA ACUACAAGCUGCCCGACGACUUCACCGGCUGCGUGAUCGCCUGGAA CAGCAACAACCUCGACAGCAAGGUGGGCGGCAACUACAACUACCUG UACCGGCUGUUCCGGAAGAGCAACCUGAAGCCCUUCGAGCGGGACA UCAGCACCGAGAUCUACCAAGCCGGCUCCACCCCUUGCAACGGCGU GGAGGGCUUCAACUGCUACUUCCCUCUGCAGAGCUACGGCUUCCAG CCCACCAACGGCGUGGGCUACCAGCCCUACCGGGUGGUGGUGCUGA GCUUCGAGCUGCUGCACGCCCCAGCCACCGUGUGUGGCCCCAAGAA GAGCACCAACCUGGUGAAGAACAAGUGCGUGAACUUCAACUUCAAC GGCCUUACCGGCACCGGCGUGCUGACCGAGAGCAACAAGAAAUUCC UGCCCUUUUGCCAGUUCGGCCGGGACAUCGCCGACACCACCGACGC UGUGCGGGAUCCCCAGACCCUGGAGAUCCUGGACAUCACCCCUUGC AGCUUCGGCGGCUCUGGCGGAGGCAGCAUCCUGGCCAUCUACAGCA CCGUGGCCAGCAGCCUGGUGCUGCUGGUGAGCCUGGGCGCCAUCAG CUUC	28
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGG CCUCCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGC	4
Corresponding amino acid sequence	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVCRSS VLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYF ASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCND PFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQG NFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLP IGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFL LKYNENGTITDAVDCALDPLSETKCTLKSFTVGSGGGSGGGSGGGS GGGSGGQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISN CVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDE VRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYL	29

	YRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQ PTNGVGYQPYRVVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFN GLTGTGVLTESNKKFLPFCQFGRDIADTTDAVRDPQTLEILDITPC SFGGsgggsilaiystvasslvllvslgaisf	
PolyA tail	100 nt	
SARS-CoV-2 S1 Subunit Truncated and Linked to Transmembrane Domain (S1-666-TM)		
SEQ ID NO: 149 consists of from 5' end to 3' end: 5' UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 150 and 3' UTR SEQ ID NO: 4.		149
Chemistry	1-methylpseudouridine	
Cap	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCG GCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGUGAGCAGCCAGUGCGU GAACCUGACCACCCGGACCCAGCUGCCACCAGCCUACACCAACAGCU UCACCCGGGGCGUCUACUACCCCGACAAGGUGUUCCGGAGCAGCGUC CUGCACAGCACCCAGGACCUGUUCCUGCCCUUCUUCAGCAACGUGAC CUGGUUCCACGCCAUCCACGUGAGCGGCACCAACGGCACCAAGCGGU UCGACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUCGCCAGC ACCGAGAAGAGCAACAUCAUCCGGGGCUGGAUCUUCGGCACCACCCU GGACAGCAAGACCCAGAGCCUGCUGAUCGUGAAUAACGCCACCAACG UGGUGAUCAAGGUGUGCGAGUUCCAGUUCUGCAACGACCCCUUCCUG GGCGUGUACUACCACAAGAACAACAAGAGCUGGAUGGAGAGCGAGUU CCGGGUGUACAGCAGCGCCAACAACUGCACCUUCGAGUACGUGAGCC AGCCCUUCCUGAUGGACCUGGAGGGCAAGCAGGGCAACUUCAAGAAC CUGCGGGAGUUCGUGUUCAAGAACAUCGACGGCUACUUCAAGAUCUA CAGCAAGCACACCCCAAUCAACCUGGUGCGGGAUCUGCCCCAGGGCU UCUCAGCCCUGGAGCCCCUGGUGGACCUGCCCAUCGGCAUCAACAUC ACCCGGUUCCAGACCCUGCUGGCCCUGCACCGGAGCUACCUGACCCC AGGCGACAGCAGCAGCGGGUGGACAGCAGGCGCGGCUGCUUACUACG UGGGCUACCUGCAGCCCCGGACCUUCCUGCUGAAGUACAACGAGAAC GGCACCAUCACCGACGCCGUGGACUGCGCCCUGGACCCUCUGAGCGA GACCAAGUGCACCCUGAAGAGCUUCACCGUGGAGAAGGGCAUCUACC AGACCAGCAACUUCCGGGUGCAGCCCACCGAGAGCAUCGUGCGGUUC CCCAACAUCACCAACCUGUGCCCCUUCGGCGAGGUGUUCAACGCCAC CCGGUUCGCCAGCGUGUACGCCUGGAACCGGAAGCGGAUCAGCAACU GCGUGGCCGACUACAGCGUGCUGUACAACAGCGCCAGCUUCAGCACC UUCAAGUGCUACGGCGUGAGCCCCACCAAGCUGAACGACCUGUGCUU CACCAACGUGUACGCCGACAGCUUCGUGAUCCGUGGCGACGAGGUGC GGCAGAUCGCACCCGGCCAGACAGGCAAGAUCGCCGACUACAACUAC AAGCUGCCCGACGACUUCACCGGCUGCGUGAUCGCCUGGAACAGCAA CAACCUCGACAGCAAGGUGGGCGGCAACUACAACUACCUGUACCGGC UGUUCCGGAAGAGCAACCUGAAGCCCUUCGAGCGGGACAUCAGCACC GAGAUCUACCAAGCCGGCUCCACCCCUUGCAACGGCGUGGAGGGCUU CAACUGCUACUUCCCUCUGCAGAGCUACGGCUUCCAGCCCACCAACG GCGUGGGCUACCAGCCCUACCGGGUGGUGGUGCUGAGCUUCGAGCUG CUGCACGCCCCAGCCACCGUGUGUGGCCCCAAGAAGAGCACCAACCU GGUGAAGAACAAGUGCGUGAACUUCAACUUCAACGGCCUUACCGGCA CCGGCGUGCUGACCGAGAGCAACAAGAAAUUCCUGCCCUUUCAGCAG UUCGGCCGGGACAUCGCCGACACCACCGACGCUGUGCGGGAUCCCCA GACCCUGGAGAUCCUGGACAUCACCCCUUGCAGCUUCGGCGGCGUGA GCGUGAUCACCCCAGGCACCAACACCAGCAACCAGGUGGCCGUGCUG UACCAGGACGUGAACUGCACCGAGGUGCCCGUGGCCAUCCACGCCGA CCAGCUGACACCCACCUGGCGGGUCUACAGCACCGGCAGCAACGUGU UCCAGACCCGGGCCGGUUGCCUGAUCGGCGCCGAGCACGUGAACAAC AGCUACGAGUGCGACAUCCCCAUCGGCGCCGGCAUCUGUGCCAGCUA CCAGACCCAGACCAAUUCA	150
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGG CCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGC	4

Corresponding amino acid sequence	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSV LHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFAS TEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFL GVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKN LREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINI TRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNEN GTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQPTESIVRF PNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFST FKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNY KLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDIST EIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFEL LHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQ FGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVL YQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNN SYECDIPIGAGICASYQTQTNS	151
PolyA tail	100 nt	

Table 4A. Subunit Antigen RBD Deletions

Name	SEQ ID NO:	
	mRNA ORF	Protein
SARS-CoV-2 S1 Subunit with Deletion of RBD and Glycan Added to S2 Subunit	31	32
SARS-CoV-2 S1 Subunit with Deletion of RBD and NTD and Glycan Added to S2 Subunit	34	35

Table 4B. Subunit Antigen RBD Deletions

SARS-CoV-2 S1 Subunit with Deletion of RBD and Glycan Added to S2 Subunit	
SEQ ID NO: 30 consists of from 5' end to 3' end: 5' UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 31 and 3' UTR SEQ ID NO: 4.	30
Chemistry	1-methylpseudouridine
Cap	7mG(5')ppp(5')NlmpNp
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCG GCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGUGAGCAGCCAGUGCG UGAACCUGACCACCCGGACCCAGCUGCCACCAGCCUACACCAACAG CUUCACCCGGGGCGUCUACUACCCCGACAAGGUGUUCCGGAGCAGC GUCCUGCACAGCACCCAGGACCUGUUCCUGCCCUUCUUCAGCAACG UGACCUGGUUCCACGCCAUCCACGUGAGCGGCACCAACGGCACCAA GCGGUUCGACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUC GCCAGCACCGAGAAGAGCAACAUCAUCCGGGGCUGGAUCUUCGGCA CCACCCUGGACAGCAAGACCCAGAGCCUGCUGAUCGUGAAUAACGC CACCAACGUGGUGAUCAAGGUGUGCGAGUUCCAGUUCUGCAACGAC CCCUUCCUGGGCGUGUACUACCACAAGAACAACAAGAGCUGGAUGG AGAGCGAGUUCCGGGUGUACAGCAGCGCCAACAACUGCACCUUCGA GUACGUGAGCCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAGGGC AACUUCAAGAACCUGCGGGAGUUCGUGUUCAAGAACAUCGACGGCU ACUUCAAGAUCUACAGCAAGCACACCCCAAUCAACCUGGUGCGGGA UCUGCCCCAGGGCUUCUCAGCCCUGGAGCCCCUGGUGGACCUGCCC AUCGGCAUCAACAUCACCCGGUUCCAGACCCUGCUGGCCCUGCACC GGAGCUACCUGACCCCAGGCGACAGCAGCAGCGGGUGGACAGCAGG CGCGGCUGCUUACUACGUGGGCUACCUGCAGCCCCGGACCUUCCUG CUGAAGUACAACGAGAACGGCACCAUCACCGACGCCGUGGACUGCG CCCUGGACCCUCUGAGCGAGACCAAGUGCACCCUGAAGAGCUUCAC CGUGGAGAAGGGCAUCUACCAGACCAGCAACUUCGGCGGCAGCGGC GGCGUGAGCGUGAUCACCCCAGGCACCAACACCAGCAACCAGGUGG	31

	CCGUGCUGUACCAGGACGUGAACUGCACCGAGGUGCCCGUGGCCAU CCACGCCGACCAGCUGACACCCACCUGGCGGGUCUACAGCACCGGC AGCAACGUGUUCCAGACCCGGGCCGGUUGCCUGAUCGGCGCCGAGC ACGUGAACAACAGCUACGAGUGCGACAUCCCCAUCGGCGCCGGCAU CUGUGCCAGCUACCAGACCCAGACCAAUUCACCCCGGAGGGCAAGG AGCGUGGCCAGCCAGAGCAUCAUCGCCUACACCAUGAGCCUGGGCG CCGAGAACAGCGUGGCCUACAGCAACAACAGCAUCGCCAUCCCCAC CAACUUCACCAUCAGCGUGACCACCGAGAUUCUGCCCGUGAGCAUG ACCAAGACCAGCGUGGACUGCACCAUGUACAUCUGCGGCGACAGCA CCGAGUGCAGCAACCUGCUGCUGCAGUACGGCAGCUUCUGCACCCA GCUGAACCGGGCCCUGACCGGCAUCGCCGUGGAGCAGGACAAGAAC ACCCAGGAGGUGUUCGCCCAGGUGAAGCAGAUCUACAAGACCCCUC CCAUCAAGGACUUCGGCGGCUUCAACUUCAGCCAGAUCCUGCCCGA CCCCAGCAAGCCCAGCAAGCGGAGCUUCAUCGAGGACCUGCUGUUC AACAAGGUGACCCUAGCCGACGCCGGCUUCAUCAAGCAGUACGGCG ACUGCCUCGGCGACAUAGCCGCCCGGGACCUGAUCUGCGCCCAGAA GUUCAACGGCCUGACCGUGCUGCCUCCCCUGCUGACCGACGAGAUG AUCGCCCAGUACACCAGCGCCCUGUUAGCCGGAACCAUCACCAGCG GCUGGACUUUCGGCGCUGGAGCCGCUCUGCAGAUCCCCUUCGCCAU GCAGAUGGCCUACCGGUUCAACGGCAUCGGCGUGACCCAGAACGUG CUGUACGAGAACCAGAAGCUGAUCGCCAACCAGUUCAACAGCGCCA UCGGCAAGAUCCAGGACAGCCUGAGCAGCACCGCUAGCGCCCUGGG CAAGCUGCAGGACGUGGUGAACCAGAACGCCCAGGCCCUGAACACC CUGGUGAAGCAGCUGAGCAGCAACUUCGGCGCCAUCAGCAGCGUGC UGAACGACAUCCUGAGCCGGCUGGACCCUCCCAACGCCACCGUGCA GAUCGACCGGCUGAUCACUGGCCGGCUGCAGAGCCUGCAGACCUAC GUGACCCAGCAGCUGAUCCGGGCCGCCGAGAUUCGGGCCAGCGCCA ACCUGGCCGCCACCAAGAUGAGCGAGUGCGUGCUGGGCCAGAGCAA GCGGGUGGACUUCUGCGGCAAGGGCUACCACCUGAUGAGCUUUCCC CAGAGCGCACCCCACGGAGUGGUGUUCCUGCACGUGACCUACGUGC CCGCCCAGGAGAAGAACUUCACCACCGCCCCAGCCAUCUGCCACGA CGGCAAGGCCCACUUUCCCCGGGAGGGCGUGUUCGUGAGCAACGGC ACCCACUGGUUCGUGACCCAGCGGAACUUCUACGAGCCCCAGAUCA UCACCACCGACAACACCUUCGUGAGCGGCAACUGCGACGUGGUGAU CGGCAUCGUGAACAACACCGUGUACGAUCCCCUGCAGCCCGAGCUG GACAGCUUCAAGGAGGAGCUGGACAAGUACUUCAAGAAUCACACCA GCCCCGACGUGGACCUGGGCGACAUCAGCGGCAUCAACGCCAGCGU GGUGAACAUCCAGAAGGAGAUCGAUCGGCUGAACGAGGUGGCCAAG AACCUGAACGAGAGCCUGAUCGACCUGCAGGAGCUGGGCAAGUACG AGCAGUACAUCAAGUGGCCCUGGUACAUCUGGCUGGGCUUCAUCGC CGGCCUGAUCGCCAUCGUGAUGGUGACCAUCAUGCUGUGCUGCAUG ACCAGCUGCUGCAGCUGCCUGAAGGGCUGUUGCAGCUGCGGCAGCU GCUGCAAGUUCGACGAGGACGACAGCGAGCCCGUGCUGAAGGGCGU GAAGCUGCACUACACC	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGG CCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGC	4
Corresponding amino acid sequence	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSS VLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYF ASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCND PFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQG NFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLP IGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFL LKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNFGGSG GVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTG SNVFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPRRAR SVASQSIIAYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSM TKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKN TQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLF NKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPLLTDEM IAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNV LYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNQNAQALNT	32

	LVKQLSSNFGAISSVLNDILSRLDPPNATVQIDRLITGRLQSLQTY VTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFP QSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFVSNG THWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPEL DSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVAK NLNESLIDLQELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCM TSCCSCLKGCCSCGSCCKFDEDDSEPVLKGVKLHYT	
PolyA tail	100 nt	
SARS-CoV-2 S1 Subunit with Deletion of RBD and NTD and Glycan Added to S2 Subunit		
SEQ ID NO: 33 consists of from 5' end to 3' end: 5' UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 34 and 3' UTR SEQ ID NO: 4.		33
Chemistry	1-methylpseudouridine	
Cap	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCG GCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGUGAGCAGCCAGGGCA CCAUCACCGACGCCGUGGACUGCGCCCUGGACCCUCUGAGCGAGAC CAAGUGCACCCUGAAGAGCUUCACCGUGGAGAAGGGCAUCUACCAG ACCAGCAACUUCGGCGGCAGCGGCGGCGUGAGCGUGAUCACCCCAG GCACCAACACCAGCAACCAGGUGGCCGUGCUGUACCAGGACGUGAA CUGCACCGAGGUGCCCGUGGCCAUCCACGCCGACCAGCUGACACCC ACCUGGCGGGUCUACAGCACCGGCAGCAACGUGUUCCAGACCCGGG CCGGUUGCCUGAUCGGCGCCGAGCACGUGAACAACAGCUACGAGUG CGACAUCCCCAUCGGCGCCGGCAUCUGUGCCAGCUACCAGACCCAG ACCAAUUCACCCCGGAGGGCAAGGAGCGUGGCCAGCCAGAGCAUCA UCGCCUACACCAUGAGCCUGGGCGCCGAGAACAGCGUGGCCUACAG CAACAACAGCAUCGCCAUCCCCACCAACUUCACCAUCAGCGUGACC ACCGAGAUUCUGCCCGUGAGCAUGACCAAGACCAGCGUGGACUGCA CCAUGUACAUCUGCGGCGACAGCACCGAGUGCAGCAACCUGCUGCU GAACUACACCAGCUUCUGCACCCAGCUGAACCGGGCCCUGACCGGC AUCGCCGUGGAGCAGGACAAGAACACCCAGGAGGUGUUCGCCCAGG UGAAGCAGAUCUACAAGACCCCUCCCAUCAAGGACUUCGGCGGCUU CAACUUCAGCCAGAUCCUGCCCGACCCCAGCAAGCCCAGCAAGCGG AGCUUCAUCGAGGACCUGCUGUUCAACAAGGUGACCCUAGCCGACG CCGGCUUCAUCAAGCAGUACGGCGACUGCCUCGGCGACAUAGCCGC CCGGGACCUGAUCUGCGCCCAGAAGUUCAACGGCCUGACCGUGCUG CCUCCCCUGCUGACCGACGAGAUGAUCGCCCAGUACACCAGCGCCC UGUUAGCCGGAACCAUCACCAGCGGCUGGACUUUCGGCGCUGGAGC CGCUCUGCAGAUCCCCUUCGCCAUGCAGAUGGCCUACCGGUUCAAC GGCAUCGGCGUGACCCAGAACGUGCUGUACGAGAACCAGAAGCUGA UCGCCAACCAGUUCAACAGCGCCAUCGGCAAGAUCCAGGACAGCCU GAGCAGCACCGCUAGCGCCCUGGGCAAGCUGCAGGACGUGGUGAAC CAGAACGCCCAGGCCCUGAACACCCUGGUGAAGCAGCUGAGCAGCA ACUUCGGCGCCAUCAGCAGCGUGCUGAACGACAUCCUGAGCCGGCU GGACCCUCCCAACGCCACCGUGCAGAUCGACCGGCUGAUCACUGGC CGGCUGCAGAGCCUGCAGACCUACGUGACCCAGCAGCUGAUCCGGG CCGCCGAGAUUCGGGCCAGCGCCAACCUGGCCGCCACCAAGAUGAG CGAGUGCGUGCUGGGCCAGAGCAAGCGGGUGGACUUCUGCGGCAAG GGCUACCACCUGAUGAGCUUUCCCCAGAGCGCACCCCACGGAGUGG UGUUCCUGCACGUGACCUACGUGCCCGCCCAGGAGAAGAACUUCAC CACCGCCCCAGCCAUCUGCCACGACGGCAAGGCCCACUUUCCCCGG GAGGGCGUGUUCGUGAGCAACGGCACCCACUGGUUCGUGACCCAGC GGAACUUCUACGAGCCCCAGAUCAUCACCACCGACAACACCUUCGU GAGCGGCAACUGCGACGUGGUGAUCGGCAUCGUGAACAACACCGUG UACGAUCCCCUGCAGCCCGAGCUGGACAGCUUCAAGGAGGAGCUGG ACAAGUACUUCAAGAAUCACACCAGCCCCGACGUGGACCUGGGCGA CAUCAGCGGCAUCAACGCCAGCGUGGUGAACAUCCAGAAGGAGAUC GAUCGGCUGAACGAGGUGGCCAAGAACCUGAACGAGAGCCUGAUCG ACCUGCAGGAGCUGGGCAAGUACGAGCAGUACAUCAAGUGGCCCUG GUACAUCUGGCUGGGCUUCAUCGCCGGCCUGAUCGCCAUCGUGAUG	34

	GUGACCAUCAUGCUGUGCUGCAUGACCAGCUGCUGCAGCUGCCUGA AGGGCUGUUGCAGCUGCGGCAGCUGCUGCAAGUUCGACGAGGACGA CAGCGAGCCCGUGCUGAAGGGCGUGAAGCUGCACUACACC	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGG CCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGC	4
Corresponding amino acid sequence	MFVFLVLLPLVSSQGTITDAVDCALDPLSETKCTLKSFTVEKGIYQ TSNFGGSGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTP TWRVYSTGSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQ TNSPRRARSVASQSIIAYTMSLGAENSVAYSNNSIAIPTNFTISVT TEILPVSMTKTSVDCTMYICGDSTECSNLLLNYTSFCTQLNRALTG IAVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKR SFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVL PPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFN GIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVN QNAQALNTLVKQLSSNFGAISSVLNDILSRLDPPNATVQIDRLITG RLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGK GYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPR EGVFVSNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTV YDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEI DRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGLIAIVM VTIMLCCMTSCCSCLKGCCSCGSCCKFDEDDSEPVLKGVKLHYT	35
PolyA tail	100 nt	

Chimeric S1-S2 Subunit Antigens

In some embodiments, a composition comprises an mRNA that encodes a chimeric protein, for example a chimeric S1-S2 protein with an S1 subunit from an S protein of one virus and an S2 subunit from an S protein of another, different virus. For example, an S2 subunit may be from SARS-CoV-2, while the S1 subunit may be from HKU1. As another example, an S2 subunit may be from SARS-CoV-2, while the S1 subunit may be from OC43. These chimeric proteins are likely to be opsonized by circulating antibodies specific to the S1 subunit of HKU1 or OC43 generated by previous exposures, promoting efficient uptake and cross-presentation of SARS-CoV-2 S2 subunit peptides to $CD4^+$ T cells by macrophages and dendritic cells. Opsonization by circulating antibodies also promotes capture by follicular dendritic cells for presentation to B cells with receptors specific to SARS-CoV-2 S2 subunit epitopes. Non-limiting examples of chimeric S1/S2 subunit constructs and the mRNA encoding them are provided in **Tables 5A** and **5B** below.

Table 5A. Chimeric S1 Subunit-S2 Subunit Antigens

Name	SEQ ID NO:	
	mRNA ORF	Protein
SARS-CoV-2 S2 Subunit Linked to HKU1 S1 Subunit	37	38
SARS-CoV-2 S2 Subunit Linked to OC43 S1 Subunit	40	41

Table 5B. Chimeric S1 Subunit-S2 Subunit Antigens

SARS-CoV-2 S2 Subunit Linked to HKU1 S1 Subunit		
SEQ ID NO: 36 consists of from 5' end to 3' end: 5' UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 37 and 3' UTR SEQ ID NO: 4.		36
Chemistry	1-methylpseudouridine	
Cap	7mG(5')ppp(5')NImpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCG GCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUUCCUGAUCAUCUUCAUCCUGCCCACCACCCUGGCCGUGAUCG GCGACUUCAACUGCACCAACAGCUUCAUCAACGACUACAACAAGAC CAUCCCUCGGAUCAGCGAAGACGUGGUGGACGUGUCCCUGGGCCUG GGCACCUACUACGUGCUGAACCGGGUGUACCUGAACACCACACUGC UGUUCACCGGCUACUUCCCCAAGAGCGGCGCCAACUUCCGGGACCU GGCCCUGAAGGGCAGCAUCUACCUGAGCACCUUGUGGUACAAGCCU CCCUUCCUGAGCGACUUCAAUAACGGCAUCUUCUCUAAGGUGAAGA ACACCAAGCUGUACGUAAACAACACCCUGUACAGCGAGUUCAGCAC CAUCGUGAUCGGCAGCGUGUUCGUCAACACCAGCUACACCAUCGUG GUGCAGCCCCACAACGGCAUCCUGGAGAUCACCGCCUGCCAGUACA CCAUGUGCGAGUACCCUCACACCGUGUGCAAGAGCAAGGGCUCCAU CCGGAACGAGAGCUGGCACAUCGACAGCAGCGAGCCGCUGUGCCUG UUCAAGAAGAACUUCACCUACAACGUGAGCGCCGACUGGCUGUACU UCCACUUCUACCAGGAGCGGGGCGUGUUCUACGCCUACUACGCCGA CGUGGGCAUGCCAACCACCUUCCUGUUCAGCCUGUACCUGGGCACC AUCCUGAGCCACUACUACGUGAUGCCCCUGACCUGCAACGCCAUCA GCUCAAACACCGACAACGAGACCCUGGAGUACUGGGUGACUCCACU GAGCCGGCGGCAGUACCUGCUGAACUUCGACGAGCACGGCGUGAUC ACCAACGCCGUGGACUGCGCCCUGGACCCUCUGAGCGAGACCAAGU GCACCCUGAAGAGCUUCACCGUGGAGAAGGGCAUCUACCAGACCAG CGGCUUCACCGUGAAGCCCGUAGCCACCGUGUACCGGCGGAUCCCC AACCUGCCCGACUGCGACAUCGACAACUGGCUGAACAACGUCAGCG UGCCCAGCCCACUGAACUGGGAGCGGCGGAUCUUCAGCAACUGCAA CUUCAAUCUGAGCACCCUGCUGCGGCUGGUGCACGUGGACAGCUUC AGCUGCAACAACCUGGACAAGAGCAAGAUCUUCGGUAGCUGCUUCA ACAGCAUCACCGUGGACAAGUUCGCCAUCCCUAACCGGCGGCGGGA CGAUCUGCAGCUGGGCAGCAGCGGCUUCCUGCAGAGCAGCAACUAC AAGAUCGACAUCAGCAGCUCAAGCUGCCAGCUGUACUACAGCCUGC CCCUGGUGAACGUGACCAUCAACAACUUCAACCCCAGCAGCUGGAA CCGGCGGUACGGCUUCGGCAGCUUCAACCUGAGCAGCUACGACGUG GUGUACAGCGACCACUGCUUCAGCGUGAACAGCGACUUCUGCCCCU GUGCCGACCCUAGCGUGGUGAACAGCUGCGCCAAGAGCAAGCCUCC CAGCGCCAUUUGCCCCGCCGGCACCAAGUACCGGCACUGCGACCUG GACACCACCCUGUACGUGAAGAACUGGUGCCGGUGCAGCUGCCUGC CCGACCCCAUCAGCACCUACAGCCCCAACACCUGUCCCCAGAAGAA GGUGGUGGUGGGUAUCGGCGAGCACUGUCCCGGCCUGGGCAUCAAC GAGGAGAAGUGCGGCACCCAGCUGAACCACAGCAGCUGCUUCUGUA GCCCCGACGCCUUCCUGGGCUGGAGCUUCGACAGCUGCAUCAGCAA CAACCGGUGCAACAUCUUUAGCAACUUCAUCUUCAACGGAAUCAAC AGCGGCACCACCUGCAGCAACGACCUGCUGUAUAGCAACACCGAGA UCAGCACCGGCGUGUGCGUGAACUACGACCUGUACGGCAUCACCGG CCAGGGCAUCUUCAAGGAGGUGAGCGCCGCCUACUACAACAACUGG CAGAACCUGCUGUACGACAGCAACGGCAACAUCAUCGGCUUCAAGG ACUUUCUGACCAACAAGACCUACACCAUCCUGCCCUGCUACAGCGG CGGCGUGAGCGUGAUCACCCCAGGCACCAACACCAGCAACCAGGUG GCCGUGCUGUACCAGGACGUGAACUGCACCGAGGUGCCCGUGGCCA UCCACGCCGACCAGCUGACACCCACCUGGCGGGUCUACAGCACCGG CAGCAACGUGUUCCAGACCCGGGCCGGUUGCCUGAUCGGCGCCGAG CACGUGAACAACAGCUACGAGUGCGACAUCCCCAUCGGCGCCGGCA UCUGUGCCAGCUACCAGACCCAGACCAAUUCACCCCGGAGGGCAAG GAGCGUGGCCAGCCAGAGCAUCAUCGCCUACACCAUGAGCCUGGGC GCCGAGAACAGCGUGGCCUACAGCAACAACAGCAUCGCCAUCCCCA	37

	CCAACUUCACCAUCAGCGUGACCACCGAGAUUCUGCCCGUGAGCAU GACCAAGACCAGCGUGGACUGCACCAUGUACAUCUGCGGCGACAGC ACCGAGUGCAGCAACCUGCUGCUGCAGUACGGCAGCUUCUGCACCC AGCUGAACCGGGCCCUGACCGGCAUCGCCGUGGAGCAGGACAAGAA CACCCAGGAGGUGUUCGCCCAGGUGAAGCAGAUCUACAAGACCCCU CCCAUCAAGGACUUCGGCGGCUUCAACUUCAGCCAGAUCCUGCCCG ACCCCAGCAAGCCCAGCAAGCGGAGCUUCAUCGAGGACCUGCUGUU CAACAAGGUGACCCUAGCCGACGCCGGCUUCAUCAAGCAGUACGGC GACUGCCUCGGCGACAUAGCCGCCCGGGACCUGAUCUGCGCCCAGA AGUUCAACGGCCUGACCGUGCUGCCUCCCCUGCUGACCGACGAGAU GAUCGCCCAGUACACCAGCGCCCUGUUAGCCGGAACCAUCACCAGC GGCUGGACUUUCGGCGCUGGAGCCGCUCUGCAGAUCCCCUUCGCCA UGCAGAUGGCCUACCGGUUCAACGGCAUCGGCGUGACCCAGAACGU GCUGUACGAGAACCAGAAGCUGAUCGCCAACCAGUUCAACAGCGCC AUCGGCAAGAUCCAGGACAGCCUGAGCAGCACCGCUAGCGCCCUGG GCAAGCUGCAGGACGUGGUGAACCAGAACGCCCAGGCCCUGAACAC CCUGGUGAAGCAGCUGAGCAGCAACUUCGGCGCCAUCAGCAGCGUG CUGAACGACAUCCUGAGCCGGCUGGACCCUCCCGAGGCCGAGGUGC AGAUCGACCGGCUGAUCACUGGCCGGCUGCAGAGCCUGCAGACCUA CGUGACCCAGCAGCUGAUCCGGGCCGCCGAGAUUCGGGCCAGCGCC AACCUGGCCGCCACCAAGAUGAGCGAGUGCGUGCUGGGCCAGAGCA AGCGGGUGGACUUCUGCGGCAAGGGCUACCACCUGAUGAGCUUUCC CCAGAGCGCACCCCACGGAGUGGUGUUCCUGCACGUGACCUACGUG CCCGCCCAGGAGAAGAACUUCACCACCGCCCCAGCCAUCUGCCACG ACGGCAAGGCCCACUUUCCCCGGGAGGGCGUGUUCGUGAGCAACGG CACCCACUGGUUCGUGACCCAGCGGAACUUCUACGAGCCCCAGAUC AUCACCACCGACAACACCUUCGUGAGCGGCAACUGCGACGUGGUGA UCGGCAUCGUGAACAACACCGUGUACGAUCCCCUGCAGCCCGAGCU GGACAGCUUCAAGGAGGAGCUGGACAAGUACUUCAAGAAUCACACC AGCCCCGACGUGGACCUGGGCGACAUCAGCGGCAUCAACGCCAGCG UGGUGAACAUCCAGAAGGAGAUCGAUCGGCUGAACGAGGUGGCCAA GAACCUGAACGAGAGCCUGAUCGACCUGCAGGAGCUGGGCAAGUAC GAGCAGUACAUCAAGUGGCCCUGGUACAUCUGGCUGGGCUUCAUCG CCGGCCUGAUCGCCAUCGUGAUGGUGACCAUCAUGCUGUGCUGCAU GACCAGCUGCUGCAGCUGCCUGAAGGGCUGUUGCAGCUGCGGCAGC UGCUGCAAGUUCGACGAGGACGACAGCGAGCCCGUGCUGAAGGGCG UGAAGCUGCACUACACC	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGG CCUCCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGC	4
Corresponding amino acid sequence	MFLIIFILPTTLAVIGDFNCTNSFINDYNKTIPRISEDVVDVSLGL GTYYVLNRVYLNTTLLFTGYFPKSGANFRDLALKGSIYLSTLWYKP PFLSDFNNGIFSKVKNTKLYVNNTLYSEFSTIVIGSVFVNTSYTIV VQPHNGILEITACQYTMCEYPHTVCKSKGSIRNESWHIDSSEPLCL FKKNFTYNVSADWLYFHFYQERGVFYAYYADVGMPTTFLFSLYLGT ILSHYYVMPLTCNAISSNTDNETLEYWVTPLSRRQYLLNFDEHGVI TNAVDCALDPLSETKCTLKSFTVEKGIYQTSGFTVKPVATVYRRIP NLPDCDIDNWLNNVSVPSPLNWERRIFSNCNFNLSTLLRLVHVDSF SCNNLDKSKIFGSCFNSITVDKFAIPNRRRDDLQLGSSGFLQSSNY KIDISSSSCQLYYSLPLVNVTINNFNPSSWNRRYGFGSFNLSSYDV VYSDHCFSVNSDFCPCADPSVVNSCAKSKPPSAICPAGTKYRHCDL DTTLYVKNWCRCSCLPDPISTYSPNTCPQKKVVVGIGEHCPGLGIN EEKCGTQLNHSSCFCSPDAFLGWSFDSCISNNRCNIFSNFIFNGIN SGTTCSNDLLYSNTEISTGVCVNYDLYGITGQGIFKEVSAAYYNNW QNLLYDSNGNIIGFKDFLTNKTYTILPCYSGGVSVITPGTNTSNQV AVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAE HVNNSYECDIPIGAGICASYQTQTNSPRRARSVASQSIIAYTMSLG AENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDS TECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTP PIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYG DCLGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITS GWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSA	38

	IGKIQDSLSSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISSV LNDILSRLDPPEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASA NLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYV PAQEKNFTTAPAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQI ITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHT SPDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKY EQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCLKGCCSCGS CCKFDEDDSEPVLKGVKLHYT	
PolyA tail	100 nt	
SARS-CoV-2 S2 Subunit Linked to OC43 S1 Subunit		
SEQ ID NO: 39 consists of from 5' end to 3' end: 5' UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 40 and 3' UTR SEQ ID NO: 4.		39
Chemistry	1-methylpseudouridine	
Cap	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCG GCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUUCCUGAUCCUGCUGAUCAGCCUGCCCACCGCCUUCGCCGUGA UCGGCGACCUGAAGUGCACCAGCGACAACAUCAACGACAAGGACAC CGGCCCACCACCCAUCAGCACCGACACCGUGGACGUGACCAACGGC CUGGGCACCUACUACGUGCUGGACCGGGUGUACCUGAACACCACCC UGUUCCUGAACGGCUACUACCCCACCAGCGGCAGCACCUACCGGAA UAUGGCCCUGAAGGGCAGCGUGCUGCUGAGCCGGCUGUGGUUCAAG CCACCAUUCCUGAGCGACUUCAUCAACGGAAUCUUCGCCAAGGUGA AGAACACCAAGGUGAUCAAGGACCGGGUGAUGUACAGCGAGUUCCC CGCCAUCACCAUUGGCAGUACCUUCGUGAACACCAGCUACAGCGUG GUGGUGCAGCCCCGGACCAUCAACAGCACCCAGGACGGCGACAACA AGCUGCAGGGCCUGCUGGAGGUGAGCGUGUGCCAGUACAACAUGUG CGAGUACCCUCAGACCAUCUGCCACCCCAACCUGGGCAACCACCGG AAGGAGCUGUGGCACCUGGACACCGGCGUGGUGAGCUGCCUGUACA AGCGGAACUUCACCUACGACGUAAACGCCGACUACCUGUACUUCCA CUUCUACCAGGAGGGCGGCACCUUCUACGCCUACUUCACCGACACG GGCGUGGUGACCAAGUUCCUGUUCAACGUGUACCUGGGCAUGGCCC UGAGCCACUACUACGUGAUGCCCCUGACCUGUAACAGCAAGCUGAC CCUGGAGUACUGGGUGACCCCUCUGACCAGCCGGCAGUACCUGCUG GCCUUCAACCAGGACGGCAUCAUCUUCAACGCCGUGGACUGCGCCC UGGACCCUCUGAGCGAGACCAAGUGCACCCUGAAGAGCUUCACCGU GGAGAAGGGCAUCUACCAGACCAACGGCUACACCGUGCAGCCCAUC GCCGACGUGUACCGGCGGAAGCCCAACCUGCCCAACUGCAACAUCG AGGCCUGGCUGAACGACAAGAGCGUGCCCUCGCCCCUGAACUGGGA GCGGAAGACCUUCAGCAACUGCAACUUCAACAUGAGCAGUCUGAUG AGCUUCAUCCAGGCCGACAGCUUCACCUGCAACAACAUCGACGCCG CCAAGAUCUACGGCAUGUGCUUCAGCAGCAUCACCAUCGACAAGUU UGCCAUCCCCAACGGCCGGAAGGUGGACCUGCAGCUGGGCAACCUG GGCUACCUGCAGAGCUUCAACUACCGGAUCGACACCACCGCCACCU CUUGCCAGCUGUACUACAACCUGCCCGCCGCCAACGUGAGCGUGAG CCGGUUCAACCCCAGCACCUGGAACAAGCGGUUCGGCUUCAUUGAG GACAGCGUGUUCAAACCCCGGCCCGCAGGAGUACUGACCAACCACG ACGUGGUGUACGCCCAGCACUGCUUCAAGGCACCCAAGAACUUCUG CCCCUGCAAGCUGAACGGCAGCUGUGUGGGCUCUGGCCCCGGUAAG AACAACGGCAUAGGGACUUGCCCGGCAGGGACCAACUACCUGACCU GCGACAACCUGUGCACACCCGACCCCAUCACCUUCACCGGCACCUA CAAGUGUCCCCAGACCAAGAGCCUGGUGGGCAUCGGCGAGCACUGC AGCGGCCUGGCCGUGAAGAGCGACUACUGCGGCGGCAACAGCUGCA CCUGUCGGCCCCAGGCCUUCCUGGGCUGGAGCGCCGACAGCUGCCU GCAGGGCGACAAGUGCAACAUCUUUGCCAACUUCAUCCUGCACGAC GUGAACAGCGGCCUGACCUGCAGCACCGACCUGCAGAAGGCCAACA CCGACAUCAUCCUGGGCGUGUGCGUGAACUACGACUUGUACGGCAU CCUGGGCCAGGGCAUCUUCGUGGAGGUGAACGCCACCUACUACAAC AGCUGGCAGAACCUGCUGUACGACAGCAACGGCAACCUGUACGGCU UCCGGGACUACAUCAUCAACCGGACCUUCAUGAUCCGGAGCUGCUA	40

	CAGCGGCGGCGUGAGCGUGAUCACCCCAGGCACCAACACCAGCAAC CAGGUGGCCGUGCUGUACCAGGACGUGAACUGCACCGAGGUGCCCG UGGCCAUCCACGCCGACCAGCUGACACCCACCUGGCGGGUCUACAG CACCGGCAGCAACGUGUUCCAGACCCGGGCCGGUUGCCUGAUCGGC GCCGAGCACGUGAACAACAGCUACGAGUGCGACAUCCCCAUCGGCG CCGGCAUCUGUGCCAGCUACCAGACCCAGACCAAUUCACCCCGGAG GGCAAGGAGCGUGGCCAGCCAGAGCAUCAUCGCCUACACCAUGAGC CUGGGCGCCGAGAACAGCGUGGCCUACAGCAACAACAGCAUCGCCA UCCCCACCAACUUCACCAUCAGCGUGACCACCGAGAUUCUGCCCGU GAGCAUGACCAAGACCAGCGUGGACUGCACCAUGUACAUCUGCGGC GACAGCACCGAGUGCAGCAACCUGCUGCUGCAGUACGGCAGCUUCU GCACCCAGCUGAACCGGGCCCUGACCGGCAUCGCCGUGGAGCAGGA CAAGAACACCCAGGAGGUGUUCGCCCAGGUGAAGCAGAUCUACAAG ACCCCUCCCAUCAAGGACUUCGGCGGCUUCAACUUCAGCCAGAUCC UGCCCGACCCCAGCAAGCCCAGCAAGCGGAGCUUCAUCGAGGACCU GCUGUUCAACAAGGUGACCCUAGCCGACGCCGGCUUCAUCAAGCAG UACGGCGACUGCCUCGGCGACAUAGCCGCCCGGGACCUGAUCUGCG CCCAGAAGUUCAACGGCCUGACCGUGCUGCCUCCCCUGCUGACCGA CGAGAUGAUCGCCCAGUACACCAGCGCCCUGUUAGCCGGAACCAUC ACCAGCGGCUGGACUUUCGGCGCUGGAGCCGCUCUGCAGAUCCCCU UCGCCAUGCAGAUGGCCUACCGGUUCAACGGCAUCGGCGUGACCCA GAACGUGCUGUACGAGAACCAGAAGCUGAUCGCCAACCAGUUCAAC AGCGCCAUCGGCAAGAUCCAGGACAGCCUGAGCAGCACCGCUAGCG CCCUGGGCAAGCUGCAGGACGUGGUGAACCAGAACGCCCAGGCCCU GAACACCCUGGUGAAGCAGCUGAGCAGCAACUUCGGCGCCAUCAGC AGCGUGCUGAACGACAUCCUGAGCCGGCUGGACCCUCCCGAGGCCG AGGUGCAGAUCGACCGGCUGAUCACUGGCCGGCUGCAGAGCCUGCA GACCUACGUGACCCAGCAGCUGAUCCGGGCCGCCGAGAUUCGGGCC AGCGCCAACCUGGCCGCCACCAAGAUGAGCGAGUGCGUGCUGGGCC AGAGCAAGCGGGUGGACUUCUGCGGCAAGGGCUACCACCUGAUGAG CUUUCCCCAGAGCGCACCCCACGGAGUGGUGUUCCUGCACGUGACC UACGUGCCCGCCCAGGAGAAGAACUUCACCACCGCCCCAGCCAUCU GCCACGACGGCAAGGCCCACUUUCCCCGGGAGGGCGUGUUCGUGAG CAACGGCACCCACUGGUUCGUGACCCAGCGGAACUUCUACGAGCCC CAGAUCAUCACCACCGACAACACCUUCGUGAGCGGCAACUGCGACG UGGUGAUCGGCAUCGUGAACAACACCGUGUACGAUCCCCUGCAGCC CGAGCUGGACAGCUUCAAGGAGGAGCUGGACAAGUACUUCAAGAAU CACACCAGCCCCGACGUGGACCUGGGCGACAUCAGCGGCAUCAACG CCAGCGUGGUGAACAUCCAGAAGGAGAUCGAUCGGCUGAACGAGGU GGCCAAGAACCUGAACGAGAGCCUGAUCGACCUGCAGGAGCUGGGC AAGUACGAGCAGUACAUCAAGUGGCCCUGGUACAUCUGGCUGGGCU UCAUCGCCGGCCUGAUCGCCAUCGUGAUGGUGACCAUCAUGCUGUG CUGCAUGACCAGCUGCUGCAGCUGCCUGAAGGGCUGUUGCAGCUGC GGCAGCUGCUGCAAGUUCGACGAGGACGACAGCGAGCCCGUGCUGA AGGGCGUGAAGCUGCACUACACC	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGG CCUCCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGC	4
Corresponding amino acid sequence	MFLILLLISLPTAFAVIGDLKCTSDNINDKDTGPPPISTDTVDVTNG LGTYYVLDRVYLNTTLFLNGYYPTSGSTYRNMALKGSVLLSRLWFK PPFLSDFINGIFAKVKNTKVIKDRVMYSEFPAITIGSTFVNTSYSV VVQPRTINSTQDGDNKLQGLLEVSVCQYNMCEYPQTICHPNLGNHR KELWHLDTGVVSCLYKRNFTYDVNADYLYFHFYQEGGTFYAYFTDT GVVTKFLFNVYLGMALSHYYVMPLTCNSKLTLEYWVTPLTSRQYLL AFNQDGIIFNAVDCALDPLSETKCTLKSFTVEKGIYQTNGYTVQPI ADVYRRKPNLPNCNIEAWLNDKSVPSPLNWERKTFSNCNFNMSSLM SFIQADSFTCNNIDAAKIYGMCFSSITIDKFAIPNGRKVDLQLGNL GYLQSFNYRIDTTATSCQLYYNLPAANVSVSRFNPSTWNKRFGFIE DSVFKPRPAGVLTNHDVVYAQHCFKAPKNFCPCKLNGSCVGSGPGK NNGIGTCPAGTNYLTCDNLCTPDPITFTGTYKCPQTKSLVGIGEHC SGLAVKSDYCGGNSCTCRPQAFLGWSADSCLQGDKCNIFANFILHD VNSGLTCSTDLQKANTDIILGVCVNYDLYGILGQGIFVEVNATYYN	41

	SWQNLLYDSNGNLYGFRDYIINRTFMIRSCYSGGVSVITPGTNTSN QVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIG AEHVNNSYECDIPIGAGICASYQTQTNSPRRARSVASQSIIAYTMS LGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICG DSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYK TPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQ YGDCLGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTI TSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFN SAIGKIQDSLSSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAIS SVLNDILSRLDPPEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRA SANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVT YVPAQEKNFTTAPAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEP QIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKN HTSPDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELG KYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCLKGCCSC GSCCKFDEDDSEPVLKGVKLHYT	
PolyA tail	100 nt	

Encoded Domain Antigens

Other aspects of the present disclosure provide compositions comprising an mRNA that encodes a (at least one) subdomain of the SARS-CoV-2 S1 subunit of the S protein. The subdomain may be an N-terminal domain (NTD) or a receptor binding domain (RBD) (with or without the SD1 and/or SD2). In some embodiments, an mRNA encodes a combination (e.g., a non-natural combination) of an NTD and an RBD (with or without the SD1 and/or SD2). In some embodiments, the NTD and/or RBD is linked to a transmembrane domain (with or without the SD1 and/or SD2). In some embodiments, the mRNA encodes two subdomains of the SARS-CoV-2 S1 subunit of the S protein (NTD and RBD) that have been mutated to comprise cysteine residues. Such mutations, in some embodiments, result in the formation of a disulfide bond. As an example, an mRNA may encode an NTD comprising an F43C mutation and an RBD comprising a Q563C mutation, ultimately resulting in a an NTD linked to an RBD via disulfide bond.

N Terminal Domain (NTD) Constructs

In some embodiments, an mRNA provided herein encodes an NTD of an S1 subunit of a SARS-CoV-2 S protein. The NTD of certain betacoronaviruses elicits protective levels of antibodies. Antibodies specific to the NTD of other betacoronaviruses such as MERS act by preventing membrane fusion and viral entry (Zhou H et al. Nat Commun. 2019; 3068), providing a second mechanism of neutralization that is distinct from preventing viral attachment to ACE2. The SARS-CoV-2 NTDs encoded by an mRNA of the present disclosure may be soluble or membrane bound. A non-limiting example of a membrane bound SARS-CoV-2 NTD antigen and the mRNA encoding it is provided in **Tables 6A** and **6B** below.

Table 6A. Membrane Bound NTD Antigens

Name	SEQ ID NO:	
	mRNA ORF	Protein
SARS-CoV-2 NTD Linked to Transmembrane Domain (NTD-TM)	46	47

Table 6B. Membrane Bound NTD Antigens

SARS-CoV-2 NTD Linked to Transmembrane Domain (NTD-TM)		
SEQ ID NO: 45 consists of from 5' end to 3' end: 5' UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 46 and 3' UTR SEQ ID NO: 4.		45
Chemistry	1-methylpseudouridine	
Cap	7mG(5')ppp(5')NImpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCG GCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGUGAGCAGCCAGUGCG UGAACCUGACCACCCGGACCCAGCUGCCACCAGCCUACACCAACAG CUUCACCCGGGGCGUCUACUACCCCGACAAGGUGUUCCGGAGCAGC GUCCUGCACAGCACCCAGGACCUGUUCCUGCCCUUCUUCAGCAACG UGACCUGGUUCCACGCCAUCCACGUGAGCGGCACCAACGGCACCAA GCGGUUCGACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUC GCCAGCACCGAGAAGAGCAACAUCAUCCGGGGCUGGAUCUUCGGCA CCACCCUGGACAGCAAGACCCAGAGCCUGCUGAUCGUGAAUAACGC CACCAACGUGGUGAUCAAGGUGUGCGAGUUCCAGUUCUGCAACGAC CCCUUCCUGGGCGUGUACUACCACAAGAACAACAAGAGCUGGAUGG AGAGCGAGUUCCGGGUGUACAGCAGCGCCAACAACUGCACCUUCGA GUACGUGAGCCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAGGGC AACUUCAAGAACCUGCGGGAGUUCGUGUUCAAGAACAUCGACGGCU ACUUCAAGAUCUACAGCAAGCACACCCCAAUCAACCUGGUGCGGGA UCUGCCCCAGGGCUUCUCAGCCCUGGAGCCCCUGGUGGACCUGCCC AUCGGCAUCAACAUCACCCGGUUCCAGACCCUGCUGGCCCUGCACC GGAGCUACCUGACCCCAGGCGACAGCAGCAGCGGGUGGACAGCAGG CGCGGCUGCUUACUACGUGGGCUACCUGCAGCCCCGGACCUUCCUG CUGAAGUACAACGAGAACGGCACCAUCACCGACGCCGUGGACUCUG GCGGAGGCAGCAUCCUGGCCAUCUACAGCACCGUGGCCAGCAGCCU GGUGCUGCUGGUGAGCCUGGGCGCCAUCAGCUUC	46
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGG CCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGC	4
Corresponding amino acid sequence	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSS VLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYF ASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCND PFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQG NFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLP IGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFL LKYNENGTITDAVDSGGGSILAIYSTVASSLVLLVSLGAISF	47
PolyA tail	100 nt	

Receptor Binding Domain (RBD) Constructs

In other embodiments, an mRNA provided herein encodes an RBD of an S1 subunit of a SARS-CoV-2 S protein. The RBD binds ACE2 receptors on host cells, which mediate virus attachment to cells. Attachment is necessary for the virus to enter cells and replicate. Thus, RBD

targeted antibody responses, which block virus attachment into the cell, effectively neutralize extracellular virus particles, preventing proliferation and promoting further immune responses to other components of the neutralized virus particles. The SARS-CoV-2 RBDs encoded by an mRNA of the present disclosure may be soluble or membrane bound (e.g., linked to a transmembrane domain).

Soluble RBD Antigens

In some embodiments, an mRNA encodes a soluble SARS-CoV-2 RBD. Dendritic cells sample soluble proteins by pinocytosis and, upon migrating to the draining lymph node, present linear peptides that comprise the sampled protein to $CD4^+$ T cells. These $CD4^+$ T cells provide proliferation signals to B cells that have recognized, taken up, and presented an epitope from the RBD, so administration of specifically RBD without other components of the SARS-CoV-2 spike protein expected to focus the immune response towards the epitopes present in the RBD. Non-limiting examples of soluble SARS-CoV-2 RBDs and the mRNA encoding them are provided in the **Tables 7A** and **7B** below.

Table 7. Soluble RBD Antigens

Name	SEQ ID NO:	
	mRNA ORF	Protein
SARS-CoV-2 Soluble RBD	61	62

Table 7B. Soluble RBD Antigens

SARS-CoV-2 Soluble RBD		
SEQ ID NO: 60 consists of from 5' end to 3' end: 5' UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 61 and 3' UTR SEQ ID NO: 4.		60
Chemistry	1-methylpseudouridine	
Cap	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCG GCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUACAGCAUGCAGCUGGCUAGCUGCGUGACCCUGACCCUGGUGC UGCUGGUGAACAGCCAGCCCAACAUCACCAACCUGUGCCCCUUCGG CGAGGUGUUCAACGCCACCCGGUUCGCCAGCGUGUACGCCUGGAAC CGGAAGCGGAUCAGCAACUGCGUGGCCGACUACAGCGUGCUGUACA ACAGCGCCAGCUUCAGCACCUUCAAGUGCUACGGCGUGAGCCCCAC CAAGCUGAACGACCUGUGCUUCACCAACGUGUACGCCGACAGCUUC GUGAUCCGUGGCGACGAGGUGCGGCAGAUCGCACCCGGCCAGACAG GCAAGAUCGCCGACUACAACUACAAGCUGCCCGACGACUUCACCGG CUGCGUGAUCGCCUGGAACAGCAACAACCUCGACAGCAAGGUGGGC GGCAACUACAACUACCUGUACCGGCUGUUCCGGAAGAGCAACCUGA AGCCCUUCGAGCGGGACAUCAGCACCGAGAUCUACCAAGCCGGCUC CACCCCUUGCAACGGCGUGGAGGGCUUCAACUGCUACUUCCCUCUG CAGAGCUACGGCUUCCAGCCCACCAACGGCGUGGGCUACCAGCCCU	61

	ACCGGGUGGUGGUGCUGAGCUUCGAGCUGCUGCACGCCCCAGCCAC CGUGUGUGGCCCCAAG	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGG CCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGC	4
Corresponding amino acid sequence	MYSMQLASCVTLTLVLLVNSQPNITNLCPFGEVFNATRFASVYAWN RKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSF VIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVG GNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPL QSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPK	62
PolyA tail	100 nt	

Membrane Bound RBD Antigens

In some embodiments, an mRNA encodes a membrane bound SARS-CoV-2 RBD. Cells expressing membrane bound RBD are expected to carry these membrane-bound antigens to the draining lymph node and promote efficient recognition of epitopes by RBD-specific B cells. Because the B cell surface contains many surface bound antibodies and the expressing cell contains many copies of the membrane bound RBD, it is expected that initial recognition of antigen by a B cell will be followed by cross-linking of B cell receptors, stimulating a strong response through an avidity effect. Non-limiting examples of membrane bound SARS-CoV-2 RBDs and the mRNA encoding them are provided in **Tables 8A** and **8B** below.

Table 8A. Membrane Bound RBD Antigens

Name	SEQ ID NO:	
	mRNA ORF	Protein
SARS-CoV-2 RBD Linked to Transmembrane Domain (RBD-TM)	76	77

Table 8B. Membrane Bound RBD Antigens

SARS-CoV-2 RBD Linked to Transmembrane Domain (RBD-TM)		
SEQ ID NO: 75 consists of from 5' end to 3' end: 5' UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 76 and 3' UTR SEQ ID NO: 4.		75
Chemistry	1-methylpseudouridine	
Cap	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCG GCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUACAGCAUGCAGCUGGCUAGCUGCGUGACCCUGACCCUGGUGC UGCUGGUGAACAGCCAGCCCAACAUCACCAACCUGUGCCCCUUCGG CGAGGUGUUCAACGCCACCCGGUUCGCCAGCGUGUACGCCUGGAAC CGGAAGCGGAUCAGCAACUGCGUGGCCGACUACAGCGUGCUGUACA ACAGCGCCAGCUUCAGCACCUUCAAGUGCUACGGCGUGAGCCCCAC CAAGCUGAACGACCUGUGCUUCACCAACGUGUACGCCGACAGCUUC GUGAUCCGUGGCGACGAGGUGCGGCAGAUCGCACCCGGCCAGACAG GCAAGAUCGCCGACUACAACUACAAGCUGCCCGACGACUUCACCGG CUGCGUGAUCGCCUGGAACAGCAACAACCUCGACAGCAAGGUGGGC	76

	GGCAACUACAACUACCUGUACCGGCUGUUCCGGAAGAGCAACCUGA AGCCCUUCGAGCGGGACAUCAGCACCGAGAUCUACCAAGCCGGCUC CACCCCUUGCAACGGCGUGGAGGGCUUCAACUGCUACUUCCCUCUG CAGAGCUACGGCUUCCAGCCCACCAACGGCGUGGGCUACCAGCCCU ACCGGGUGGUGGUGCUGAGCUUCGAGCUGCUGCACGCCCCAGCCAC CGUGUGUGGCCCCAAGUCUGGCGGAGGCAGCAUCCUGGCCAUCUAC AGCACCGUGGCCAGCAGCCUGGUGCUGCUGGUGAGCCUGGGCGCCA UCAGCUUC	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGG CCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGC	4
Corresponding amino acid sequence	MYSMQLASCVTLTLVLLVNSQPNITNLCPFGEVFNATRFASVYAWN RKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSF VIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVG GNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPL QSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKSGGGSILAIY STVASSLVLLVSLGAISF	77
PolyA tail	100 nt	

Domain Fusion Antigens

In yet other embodiments, an mRNA provided herein encodes a SARS-CoV-2 NTD-RBD fusion protein. For example, the NTD and the RBD of a SARS-CoV-2 S1 subunit of an S protein may be linked to each other through a linker, such as a short amino acid (e.g., glycine-serine) linker to allow flexibility/hinging and space between the domains. In another embodiment, a linker comprising an antigenic epitope, e.g., a Class II universal T cell epitope such as PADRE, can be used. In some embodiments, a transmembrane region is linked to the NTD-RBD fusion, for example, through another short amino acid (e.g., glycine-serine or PADRE) linker for flexibility and to permit a reasonable distance between the membrane and the antigen. Without being bound by theory, it is thought that this membrane bound, tandem configuration presents most, if not all, known neutralizing and protective epitopes in one open reading frame. Administration of this fusion protein should then focus the immune response towards known protective epitopes and reduce the unnecessary generation of antibodies and T cells specific to non-protective epitopes. Furthermore, antibodies to different domains may neutralize virus particles through different mechanisms, such as by blocking attachment to host cells or preventing bound virus from undergoing membrane fusion and entering host cells. The broad response elicited by a fusion protein comprising different domains may thus be more evolutionarily robust, requiring multiple distinct mutations to escape vaccine-induced immunity. Non-limiting examples of SARS-CoV-2 NTD-RBD fusion proteins and the mRNA encoding them are provided in **Tables 9A** and **9B** below.

Linkers

A variety of linkers may be used in accordance with the present disclosure. Linkers, as provide herein, are simply amino acid sequences that artificially link together two other amino acid sequences. Linkers used herein may be cleavable or non-cleavable. Cleavable linkers allow an mRNA to be translated into a polypeptide, after which cleavage of the linker allows each individual component to be released independently. Non-cleavable linkers keep one or more protein subunits connected, allowing the whole protein to perform a function that requires close proximity of the component subunits. Non-limiting examples of such linkers include glycine-serine (GS) linkers (non-cleavable); and F2A linker, P2A linker, T2A linker, and E2Alinker (cleavable). Other links may be used herein.

In some embodiments, the linker is a GS linker. GS linkers are polypeptide linkers that include glycine and serine amino acids repeats. They comprise flexible and hydrophilic residues and can be used to perform fusion of protein subunits without interfering in the folding and function of the protein domains, and without formation of secondary structures. In some embodiments, an mRNA encodes a fusion protein that comprises a GS linker that is 3 to 20 amino acids long. For example, the GS linker may have a length of (or have a length of at least) 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 amino acids. In some embodiments, a GS linker is (or is at least) 15 amino acids long (e.g., GGSGGSGGSGGSGGG (SEQ ID NO: 133)). In some embodiments, a GS linker is (or is at least) 8 amino acids long (e.g., GGGSGGGS (SEQ ID NO: 134)). In some embodiments, a GS linker is (or is at least) 7 amino acids long (e.g., GGGSGGG (SEQ ID NO: 135)). In some embodiments, a GS linker is (or is at least) 4 amino acid long (e.g., GGGS (SEQ ID NO: 136)). In some embodiments, the GS linker comprises (GGGS)n (SEQ ID NO: 136), where n is any integer from 1-5. In some embodiments, a GS linker is (or is at least) 4 amino acid long (e.g., GSGG (SEQ ID NO: 152)). In some embodiments, the GS linker comprises (GSGG)n (SEQ ID NO: 152), where n is any integer from 1-5.

In some embodiments, a linker is a glycine linker, for example having a length of (or a length of at least) 3 amino acids (e.g., GGG).

In some embodiments, a protein encoded by an mRNA vaccine includes more than one linker, which may be the same or different from each other (e.g., GGGSGGG (SEQ ID NO: 135) and GGGS (SEQ ID NO: 136) in the same S protein construct).

In some embodiments, a linker comprises mRNA encoding a pan HLA DR-binding epitope (PADRE) (e.g., AKFVAAWTLKAAA (SEQ ID NO: 148)). PADRE is an

immunodominant helper CD4 T cell epitope and a potent immunogen (See, e.g., Alexander J. et al. J of Immuno. 164(3): 1625-33, incorporated herein by reference).

Table 9A. Domain Fusion Antigen

Name	SEQ ID NO:	
	mRNA ORF	Protein
SARS-CoV-2 NTD-RBD Linked to Transmembrane Domain (NTD-RBD-TM)	91	92
SARS-CoV-2 RBD-NTD Linked to Transmembrane Domain (RBD-NTD-TM)	139	140
SARS-CoV-2 NTD-PADRE-RBD Linked to Transmembrane Domain (NTD-PADRE-RBD-TM)	142	143

Table 9B. Domain Fusion Antigen

SARS-CoV-2 NTD-RBD Linked to Transmembrane Domain (NTD-RBD-TM)		
SEQ ID NO: 90 consists of from 5' end to 3' end: 5' UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 91 and 3' UTR SEQ ID NO: 4.		90
Chemistry	1-methylpseudouridine	
Cap	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCG GCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGUGAGCAGCCAGUGCG UGAACCUGACCACCCGGACCCAGCUGCCACCAGCCUACACCAACAG CUUCACCCGGGGCGUCUACUACCCCGACAAGGUGUUCCGGAGCAGC GUCCUGCACAGCACCCAGGACCUGUUCCUGCCCUUCUUCAGCAACG UGACCUGGUUCCACGCCAUCCACGUGAGCGGCACCAACGGCACCAA GCGGUUCGACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUC GCCAGCACCGAGAAGAGCAACAUCAUCCGGGGCUGGAUCUUCGGCA CCACCCUGGACAGCAAGACCCAGAGCCUGCUGAUCGUGAAUAACGC CACCAACGUGGUGAUCAAGGUGUGCGAGUUCCAGUUCUGCAACGAC CCCUUCCUGGGCGUGUACUACCACAAGAACAACAAGAGCUGGAUGG AGAGCGAGUUCCGGGUGUACAGCAGCGCCAACAACUGCACCUUCGA GUACGUGAGCCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAGGGC AACUUCAAGAACCUGCGGGAGUUCGUGUUCAAGAACAUCGACGGCU ACUUCAAGAUCUACAGCAAGCACACCCCAAUCAACCUGGUGCGGGA UCUGCCCCAGGGCUUCUCAGCCCUGGAGCCCCUGGUGGACCUGCCC AUCGGCAUCAACAUCACCCGGUUCCAGACCCUGCUGGCCCUGCACC GGAGCUACCUGACCCCAGGCGACAGCAGCAGCGGGUGGACAGCAGG CGCGGCUGCUUACUACGUGGGCUACCUGCAGCCCCGGACCUUCCUG CUGAAGUACAACGAGAACGGCACCAUCACCGACGCCGUGGACGGAG GCGGAUCGGGAGGCGGACCCAACAUCACCAACCUGUGCCCCUUCGG CGAGGUGUUCAACGCCACCCGGUUCGCCAGCGUGUACGCCUGGAAC CGGAAGCGGAUCAGCAACUGCGUGGCCGACUACAGCGUGCUGUACA ACAGCGCCAGCUUCAGCACCUUCAAGUGCUACGGCGUGAGCCCCAC CAAGCUGAACGACCUGUGCUUCACCAACGUGUACGCCGACAGCUUC GUGAUCCGUGGCGACGAGGUGCGGCAGAUCGCACCCGGCCAGACAG GCAAGAUCGCCGACUACAACUACAAGCUGCCCGACGACUUCACCGG CUGCGUGAUCGCCUGGAACAGCAACAACCUCGACAGCAAGGUGGGC GGCAACUACAACUACCUGUACCGGCUGUUCCGGAAGAGCAACCUGA AGCCCUUCGAGCGGGACAUCAGCACCGAGAUCUACCAAGCCGGCUC CACCCCUUGCAACGGCGUGGAGGGCUUCAACUGCUACUUCCCUCUG CAGAGCUACGGCUUCCAGCCCACCAACGGCGUGGGCUACCAGCCCU ACCGGGUGGUGGUGCUGAGCUUCGAGCUGCUGCACGCCCCAGCCAC	91

	CGUGUGUGGCCCCAAGUCUGGCGGAGGCAGCAUCCUGGCCAUCUAC AGCACCGUGGCCAGCAGCCUGGUGCUGCUGGUGAGCCUGGGCGCCA UCAGCUUC	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGG CCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGC	4
Corresponding amino acid sequence	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSS VLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYF ASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCND PFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQG NFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLP IGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFL LKYNENGTITDAVDgggsgggPNITNLCPFGEVFNATRFASVYAWN RKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSF VIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVG GNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPL QSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKsgggsilaiy stvasslvllvslgaisf	92
PolyA tail	100 nt	
SARS-CoV-2 RBD-NTD Linked to Transmembrane Domain (RBD-NTD-TM)		
SEQ ID NO: 141 consists of from 5' end to 3' end: 5' UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 139 and 3' UTR SEQ ID NO: 4.		141
Chemistry	1-methylpseudouridine	
Cap	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCG GCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUACAGCAUGCAGCUGGCUAGCUGCGUGACCCUGACCCUGGUGC UGCUGGUGAACAGCCAGCCCAACAUCACCAACCUGUGCCCCUUCGG CGAGGUGUUCAACGCCACCCGGUUCGCCAGCGUGUACGCCUGGAAC CGGAAGCGGAUCAGCAACUGCGUGGCCGACUACAGCGUGCUGUACA ACAGCGCCAGCUUCAGCACCUUCAAGUGCUACGGCGUGAGCCCCAC CAAGCUGAACGACCUGUGCUUCACCAACGUGUACGCCGACAGCUUC GUGAUCCGUGGCGACGAGGUGCGGCAGAUCGCACCCGGCCAGACAG GCAAGAUCGCCGACUACAACUACAAGCUGCCCGACGACUUCACCGG CUGCGUGAUCGCCUGGAACAGCAACAACCUCGACAGCAAGGUGGGC GGCAACUACAACUACCUGUACCGGCUGUUCCGGAAGAGCAACCUGA AGCCCUUCGAGCGGGACAUCAGCACCGAGAUCUACCAAGCCGGCUC CACCCCUUGCAACGGCGUGGAGGGCUUCAACUGCUACUUCCCUCUG CAGAGCUACGGCUUCCAGCCCACCAACGGCGUGGGCUACCAGCCCU ACCGGGUGGUGGUGCUGAGCUUCGAGCUGCUGCACGCCCCAGCCAC CGUGUGUGGCCCCAAGUCUGGCGGAGGCAGCGGAGGCGGAAGCCAG UGCGUGAACCUGACCACCCGGACCCAGCUGCCACCAGCCUACACCA ACAGCUUCACCCGGGGCGUCUACUACCCCGACAAGGUGUUCCGGAG CAGCGUCCUGCACAGCACCCAGGACCUGUUCCUGCCCUUCUUCAGC AACGUGACCUGGUUCCACGCCAUCCACGUGAGCGGCACCAACGGCA CCAAGCGGUUCGACAACCCCGUGCUGCCCUUCAACGACGGCGUGUA CUUCGCCAGCACCGAGAAGAGCAACAUCAUCCGGGGCUGGAUCUUC GGCACCACCCUGGACAGCAAGACCCAGAGCCUGCUGAUCGUGAAUA ACGCCACCAACGUGGUGAUCAAGGUGUGCGAGUUCCAGUUCUGCAA CGACCCCUUCCUGGGCGUGUACUACCACAAGAACAACAAGAGCUGG AUGGAGAGCGAGUUCCGGGUGUACAGCAGCGCCAACAACUGCACCU UCGAGUACGUGAGCCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCA GGGCAACUUCAAGAACCUGCGGGAGUUCGUGUUCAAGAACAUCGAC GGCUACUUCAAGAUCUACAGCAAGCACACCCCAAUCAACCUGGUGC GGGAUCUGCCCCAGGGCUUCUCAGCCCUGGAGCCCCUGGUGGACCU GCCCAUCGGCAUCAACAUCACCCGGUUCCAGACCCUGCUGGCCCUG CACCGGAGCUACCUGACCCCAGGCGACAGCAGCAGCGGGUGGACAG CAGGCGCGGCUGCUUACUACGUGGGCUACCUGCAGCCCCGGACCUU CCUGCUGAAGUACAACGAGAACGGCACCAUCACCGACGCCGUGGAC	139

	UCUGGCGGAGGCAGCAUCCUGGCCAUCUACAGCACCGUGGCCAGCA GCCUGGUGCUGCUGGUGAGCCUGGGCGCCAUCAGCUUC	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGG CCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGC	4
Corresponding amino acid sequence	MYSMQLASCVTLTLVLLVNSQPNITNLCPFGEVFNATRFASVYAWN RKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSF VIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVG GNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPL QSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKSGGGSGGGSQ CVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFS NVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIF GTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLGVYYHKNNKSW MESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNID GYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQTLLAL HRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVD SGGGSILAIYSTVASSLVLLVSLGAISF	140
PolyA tail	100 nt	
SARS-CoV-2 NTD-PADRE-RBD Linked to Transmembrane Domain (NTD-PADRE-RBD-TM)		
SEQ ID NO: 144 consists of from 5' end to 3' end: 5' UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 142 and 3' UTR SEQ ID NO: 4.		144
Chemistry	1-methylpseudouridine	
Cap	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCG GCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGUGAGCAGCCAGUGCG UGAACCUGACCACCCGGACCCAGCUGCCACCAGCCUACACCAACAG CUUCACCCGGGGCGUCUACUACCCCGACAAGGUGUUCCGGAGCAGC GUCCUGCACAGCACCCAGGACCUGUUCCUGCCCUUCUUCAGCAACG UGACCUGGUUCCACGCCAUCCACGUGAGCGGCACCAACGGCACCAA GCGGUUCGACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUC GCCAGCACCGAGAAGAGCAACAUCAUCCGGGGCUGGAUCUUCGGCA CCACCCUGGACAGCAAGACCCAGAGCCUGCUGAUCGUGAAUAACGC CACCAACGUGGUGAUCAAGGUGUGCGAGUUCCAGUUCUGCAACGAC CCCUUCCUGGGCGUGUACUACCACAAGAACAACAAGAGCUGGAUGG AGAGCGAGUUCCGGGUGUACAGCAGCGCCAACAACUGCACCUUCGA GUACGUGAGCCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAGGGC AACUUCAAGAACCUGCGGGAGUUCGUGUUCAAGAACAUCGACGGCU ACUUCAAGAUCUACAGCAAGCACACCCCAAUCAACCUGGUGCGGGA UCUGCCCCAGGGCUUCUCAGCCCUGGAGCCCCUGGUGGACCUGCCC AUCGGCAUCAACAUCACCCGGUUCCAGACCCUGCUGGCCCUGCACC GGAGCUACCUGACCCCAGGCGACAGCAGCAGCGGGUGGACAGCAGG CGCGGCUGCUUACUACGUGGGCUACCUGCAGCCCCGGACCUUCCUG CUGAAGUACAACGAGAACGGCACCAUCACCGACGCCGUGGACGGAG GCGCCAAGUUCGUGGCCGCCUGGACUCUGAAGGCCGCAGCCGGCGG ACCCAACAUCACCAACCUGUGCCCCUUCGGCGAGGUGUUCAACGCC ACCCGGUUCGCCAGCGUGUACGCCUGGAACCGGAAGCGGAUCAGCA ACUGCGUGGCCGACUACAGCGUGCUGUACAACAGCGCCAGCUUCAG CACCUUCAAGUGCUACGGCGUGAGCCCCACCAAGCUGAACGACCUG UGCUUCACCAACGUGUACGCCGACAGCUUCGUGAUCCGUGGCGACG AGGUGCGGCAGAUCGCACCCGGCCAGACAGGCAAGAUCGCCGACUA CAACUACAAGCUGCCCGACGACUUCACCGGCUGCGUGAUCGCCUGG AACAGCAACAACCUCGACAGCAAGGUGGGCGGCAACUACAACUACC UGUACCGGCUGUUCCGGAAGAGCAACCUGAAGCCCUUCGAGCGGGA CAUCAGCACCGAGAUCUACCAAGCCGGCUCCACCCCUUGCAACGGC GUGGAGGGCUUCAACUGCUACUUCCCUCUGCAGAGCUACGGCUUCC AGCCCACCAACGGCGUGGGCUACCAGCCCUACCGGGUGGUGGUGCU GAGCUUCGAGCUGCUGCACGCCCCAGCCACCGUGUGUGGCCCCAAG UCUGGCGGAGGCAGCAUCCUGGCCAUCUACAGCACCGUGGCCAGCA GCCUGGUGCUGCUGGUGAGCCUGGGCGCCAUCAGCUUC	142

3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGG CCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGC	4
Corresponding amino acid sequence	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSS VLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYF ASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCND PFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQG NFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLP IGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFL LKYNENGTITDAVDGGAKFVAAWTLKAAAGGPNITNLCPFGEVFNA TRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDL CFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAW NSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNG VEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPK SGGGSILAIYSTVASSLVLLVSLGAISF	143
PolyA tail	100 nt	

Trafficking Signals

In some embodiments, an mRNA encodes a SARS-CoV-2 S protein domain (e.g., NTD, RBD, or NTD-RBD fusion) linked to a Golgi trafficking signal. Non-limiting examples of such signals include macrophage markers, such as CD86 and/or CD11b, which are highly expressed and the intracellular region may control efficient export from the Golgi apparatus to the cell surface. Other cell trafficking signals (sequences) may be used herein, for example, the VSV-G cytosolic tail (VSVGct). More efficient trafficking of encoded proteins to the cell surface is expected to increase antigen availability for B cell recognition and therefore promote the generation of antibodies to the encoded SARS-CoV-2 S protein domains. Non-limiting examples of SARS-CoV-2 antigens linked to a trafficking signal and the mRNA encoding them are provided in **Tables 10A** and **10B** below.

Table 10A. Domain Fusion Antigens Linked to a Trafficking Signal

Name	SEQ ID NO:	
	mRNA ORF	Protein
SARS-CoV-2 NTD-RBD Linked to Transmembrane Domain and huCD86 (NTD-RBD-TM-CD86)	94	95
SARS-CoV-2 NTD-RBD Linked to Transmembrane Domain and huCD11B (NTD-RBD-TM-CD11B)	97	98
SARS-CoV-2 NTD-RBD Linked to VSVGct (NTD-RBD-VSVGct)	109	110

Table 10B. Domain Fusion Antigens Linked to a Trafficking Signal

SARS-CoV-2 NTD-RBD Linked to Transmembrane Domain and huCD86 (NTD-RBD-TM-CD86)		
SEQ ID NO: 93 consists of from 5' end to 3' end: 5' UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 94 and 3' UTR SEQ ID NO: 4.		93
Chemistry	1-methylpseudouridine	

Cap	7mG(5')ppp(5')NImpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCG GCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGUGAGCAGCCAGUGCG UGAACCUGACCACCCGGACCCAGCUGCCACCAGCCUACACCAACAG CUUCACCCGGGGCGUCUACUACCCCGACAAGGUGUUCCGGAGCAGC GUCCUGCACAGCACCCAGGACCUGUUCCUGCCCUUCUUCAGCAACG UGACCUGGUUCCACGCCAUCCACGUGAGCGGCACCAACGGCACCAA GCGGUUCGACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUC GCCAGCACCGAGAAGAGCAACAUCAUCCGGGGCUGGAUCUUCGGCA CCACCCUGGACAGCAAGACCCAGAGCCUGCUGAUCGUGAAUAACGC CACCAACGUGGUGAUCAAGGUGUGCGAGUUCCAGUUCUGCAACGAC CCCUUCCUGGGCGUGUACUACCACAAGAACAACAAGAGCUGGAUGG AGAGCGAGUUCCGGGUGUACAGCAGCGCCAACAACUGCACCUUCGA GUACGUGAGCCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAGGGC AACUUCAAGAACCUGCGGGAGUUCGUGUUCAAGAACAUCGACGGCU ACUUCAAGAUCUACAGCAAGCACACCCCAAUCAACCUGGUGCGGGA UCUGCCCCAGGGCUUCUCAGCCCUGGAGCCCCUGGUGGACCUGCCC AUCGGCAUCAACAUCACCCGGUUCCAGACCCUGCUGGCCCUGCACC GGAGCUACCUGACCCCAGGCGACAGCAGCAGCGGGUGGACAGCAGG CGCGGCUGCUUACUACGUGGGCUACCUGCAGCCCCGGACCUUCCUG CUGAAGUACAACGAGAACGGCACCAUCACCGACGCCGUGGACGGAG GCGGAUCGGGAGGCGGACCCAACAUCACCAACCUGUGCCCCUUCGG CGAGGUGUUCAACGCCACCCGGUUCGCCAGCGUGUACGCCUGGAAC CGGAAGCGGAUCAGCAACUGCGUGGCCGACUACAGCGUGCUGUACA ACAGCGCCAGCUUCAGCACCUUCAAGUGCUACGGCGUGAGCCCCAC CAAGCUGAACGACCUGUGCUUCACCAACGUGUACGCCGACAGCUUC GUGAUCCGUGGCGACGAGGUGCGGCAGAUCGCACCCGGCCAGACAG GCAAGAUCGCCGACUACAACUACAAGCUGCCCGACGACUUCACCGG CUGCGUGAUCGCCUGGAACAGCAACAACCUCGACAGCAAGGUGGGC GGCAACUACAACUACCUGUACCGGCUGUUCCGGAAGAGCAACCUGA AGCCCUUCGAGCGGGACAUCAGCACCGAGAUCUACCAAGCCGGCUC CACCCCUUGCAACGGCGUGGAGGGCUUCAACUGCUACUUCCCUCUG CAGAGCUACGGCUUCCAGCCCACCAACGGCGUGGGCUACCAGCCCU ACCGGGUGGUGGUGCUGAGCUUCGAGCUGCUGCACGCCCCAGCCAC CGUGUGUGGCCCCAAGUCUGGCGGAGGCAGCAUCCUGGCCAUCUAC AGCACCGUGGCCAGCAGCCUGGUGCUGCUGGUGAGCCUGGGCGCCA UCAGCUUCAAGAAGAAGAAGCGGCCACGGAACUCCUACAAGUGCGG CACCAACACCAUGGAGCGGGAGGAGAGCGAGCAGACCAAGAAGCGG GAGAAGAUCCACAUUCCUGAACGGUCCGACGAAGCCCAGCGGGUGU UCAAGAGCAGCAAGACCAGCAGCUGCGACAAGAGCGACACCUGCUU C	94
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGG CCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGC	4
Corresponding amino acid sequence	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSS VLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYF ASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCND PFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQG NFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLP IGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFL LKYNENGTITDAVDgggsgggPNITNLCPFGEVFNATRFASVYAWN RKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSF VIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVG GNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPL QSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKsgggsilaiy stvasslvllvslgaisfKKKKRPRNSYKCGTNTMEREESEQTKKR EKIHIPERSDEAQRVFKSSKTSSCDKSDTCF	95
PolyA tail	100 nt	
SARS-CoV-2 NTD-RBD Linked to Transmembrane Domain and huCD11B (NTD-RBD-TM-CD11B)		

SEQ ID NO: 96 consists of from 5' end to 3' end: 5' UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 97 and 3' UTR SEQ ID NO: 4.		96
Chemistry	1-methylpseudouridine	
Cap	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCG GCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGUGAGCAGCCAGUGCG UGAACCUGACCACCCGGACCCAGCUGCCACCAGCCUACACCAACAG CUUCACCCGGGGCGUCUACUACCCCGACAAGGUGUUCCGGAGCAGC GUCCUGCACAGCACCCAGGACCUGUUCCUGCCCUUCUUCAGCAACG UGACCUGGUUCCACGCCAUCCACGUGAGCGGCACCAACGGCACCAA GCGGUUCGACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUC GCCAGCACCGAGAAGAGCAACAUCAUCCGGGGCUGGAUCUUCGGCA CCACCCUGGACAGCAAGACCCAGAGCCUGCUGAUCGUGAAUAACGC CACCAACGUGGUGAUCAAGGUGUGCGAGUUCCAGUUCUGCAACGAC CCCUUCCUGGGCGUGUACUACCACAAGAACAACAAGAGCUGGAUGG AGAGCGAGUUCCGGGUGUACAGCAGCGCCAACAACUGCACCUUCGA GUACGUGAGCCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAGGGC AACUUCAAGAACCUGCGGGAGUUCGUGUUCAAGAACAUCGACGGCU ACUUCAAGAUCUACAGCAAGCACACCCCAAUCAACCUGGUGCGGGA UCUGCCCCAGGGCUUCUCAGCCCUGGAGCCCCUGGUGGACCUGCCC AUCGGCAUCAACAUCACCCGGUUCCAGACCCUGCUGGCCCUGCACC GGAGCUACCUGACCCCAGGCGACAGCAGCAGCGGGUGGACAGCAGG CGCGGCUGCUUACUACGUGGGCUACCUGCAGCCCCGGACCUUCCUG CUGAAGUACAACGAGAACGGCACCAUCACCGACGCCGUGGACGGAG GCGGAUCGGGAGGCGGACCCAACAUCACCAACCUGUGCCCCUUCGG CGAGGUGUUCAACGCCACCCGGUUCGCCAGCGUGUACGCCUGGAAC CGGAAGCGGAUCAGCAACUGCGUGGCCGACUACAGCGUGCUGUACA ACAGCGCCAGCUUCAGCACCUUCAAGUGCUACGGCGUGAGCCCCAC CAAGCUGAACGACCUGUGCUUCACCAACGUGUACGCCGACAGCUUC GUGAUCCGUGGCGACGAGGUGCGGCAGAUCGCACCCGGCCAGACAG GCAAGAUCGCCGACUACAACUACAAGCUGCCCGACGACUUCACCGG CUGCGUGAUCGCCUGGAACAGCAACAACCUCGACAGCAAGGUGGGC GGCAACUACAACUACCUGUACCGGCUGUUCCGGAAGAGCAACCUGA AGCCCUUCGAGCGGGACAUCAGCACCGAGAUCUACCAAGCCGGCUC CACCCCUUGCAACGGCGUGGAGGGCUUCAACUGCUACUUCCCUCUG CAGAGCUACGGCUUCCAGCCCACCAACGGCGUGGGCUACCAGCCCU ACCGGGUGGUGGUGCUGAGCUUCGAGCUGCUGCACGCCCCAGCCAC CGUGUGUGGCCCCAAGUCUGGCGGAGGCAGCAUCCUGGCCAUCUAC AGCACCGUGGCCAGCAGCCUGGUGCUGCUGGUGAGCCUGGGCGCCA UCAGCUUCAAGCGGCAGUACAAGGACAUGAUGAGCGAGGGAGGACC ACCUGGCGCUGAGCCACAG	97
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGG CCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGC	4
Corresponding amino acid sequence	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSS VLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYF ASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCND PFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQG NFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLP IGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFL LKYNENGTITDAVDgggsgggPNITNLCPFGEVFNATRFASVYAWN RKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSF VIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVG GNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPL QSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKsgggsilaiy stvasslvllvslgaisfKRQYKDMMSEGGPPGAEPQ	98
PolyA tail	100 nt	
SARS-CoV-2 NTD-RBD Linked to VSVGct (NTD-RBD-VSVGct)		

SEQ ID NO: 108 consists of from 5' end to 3' end: 5' UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 109 and 3' UTR SEQ ID NO: 4.		108
Chemistry	1-methylpseudouridine	
Cap	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCG GCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGUGAGCAGCCAGUGCG UGAACCUGACCACCCGGACCCAGCUGCCACCAGCCUACACCAACAG CUUCACCCGGGGCGUCUACUACCCCGACAAGGUGUUCCGGAGCAGC GUCCUGCACAGCACCCAGGACCUGUUCCUGCCCUUCUUCAGCAACG UGACCUGGUUCCACGCCAUCCACGUGAGCGGCACCAACGGCACCAA GCGGUUCGACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUC GCCAGCACCGAGAAGAGCAACAUCAUCCGGGGCUGGAUCUUCGGCA CCACCCUGGACAGCAAGACCCAGAGCCUGCUGAUCGUGAAUAACGC CACCAACGUGGUGAUCAAGGUGUGCGAGUUCCAGUUCUGCAACGAC CCCUUCCUGGGCGUGUACUACCACAAGAACAACAAGAGCUGGAUGG AGAGCGAGUUCCGGGUGUACAGCAGCGCCAACAACUGCACCUUCGA GUACGUGAGCCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAGGGC AACUUCAAGAACCUGCGGGAGUUCGUGUUCAAGAACAUCGACGGCU ACUUCAAGAUCUACAGCAAGCACACCCCAAUCAACCUGGUGCGGGA UCUGCCCCAGGGCUUCUCAGCCCUGGAGCCCCUGGUGGACCUGCCC AUCGGCAUCAACAUCACCCGGUUCCAGACCCUGCUGGCCCUGCACC GGAGCUACCUGACCCCAGGCGACAGCAGCAGCGGGUGGACAGCAGG CGCGGCUGCUUACUACGUGGGCUACCUGCAGCCCCGGACCUUCCUG CUGAAGUACAACGAGAACGGCACCAUCACCGACGCCGUGGACGGAG GCGGAUCGGGAGGCGGACCCAACAUCACCAACCUGUGCCCCUUCGG CGAGGUGUUCAACGCCACCCGGUUCGCCAGCGUGUACGCCUGGAAC CGGAAGCGGAUCAGCAACUGCGUGGCCGACUACAGCGUGCUGUACA ACAGCGCCAGCUUCAGCACCUUCAAGUGCUACGGCGUGAGCCCCAC CAAGCUGAACGACCUGUGCUUCACCAACGUGUACGCCGACAGCUUC GUGAUCCGUGGCGACGAGGUGCGGCAGAUCGCACCCGGCCAGACAG GCAAGAUCGCCGACUACAACUACAAGCUGCCCGACGACUUCACCGG CUGCGUGAUCGCCUGGAACAGCAACAACCUCGACAGCAAGGUGGGC GGCAACUACAACUACCUGUACCGGCUGUUCCGGAAGAGCAACCUGA AGCCCUUCGAGCGGGACAUCAGCACCGAGAUCUACCAAGCCGGCUC CACCCCUUGCAACGGCGUGGAGGGCUUCAACUGCUACUUCCCUCUG CAGAGCUACGGCUUCCAGCCCACCAACGGCGUGGGCUACCAGCCCU ACCGGGUGGUGGUGCUGAGCUUCGAGCUGCUGCACGCCCCAGCCAC CGUGUGUGGCCCCAAGUCUGGCGGAGGCAGCAGCAGCAUCGCCAGC UUCUUCUUCAUCAUCGGGCUGAUCAUCGGCCUCUUCCUGGUGCUGC GGGUGGGCAUCCACCUGUGCAUCAAGCUGAAGCACACCAAGAAGAG ACAGAUCUACACCGACAUCGAGAUGAACCGGCUGGGCAAG	109
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGG CCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGC	4
Corresponding amino acid sequence	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSS VLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYF ASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCND PFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQG NFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLP IGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFL LKYNENGTITDAVDgggsgggPNITNLCPFGEVFNATRFASVYAWN RKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSF VIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVG GNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPL QSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKsgggsSSIAS FFFIIGLIIGLFLVLRVGIHLCIKLKHTKKRQIYTDIEMNRLGK	110
PolyA tail	100 nt	

Domain Fusion C-Terminal Truncations

In other embodiments, an mRNA provided herein encodes a SARS-CoV-2 NTD-RBD fusion protein in which some portion of the C-terminal domain has been truncated/deleted. In one embodiment, 13 (or at least 13) amino acids have been deleted from the C-terminal domain of the NTD-RBD fusion protein. Deletion of these amino acids is expected to increase exposure of epitopes to antibodies, thereby stimulating a more robust immune response to protective epitopes present on the NTD and RBD domains.

A non-limiting example of SARS-CoV-2 domain fusion antigen having a C-terminal truncation and the mRNA encoding it is provided in **Tables 11A** and **11B** below.

Table 11A. Domain Fusion C-Terminal Truncation

Name	SEQ ID NO:	
	mRNA ORF	Protein
SARS-CoV-2 NTD-RBD with C-terminal Truncation of 13 Amino Acids (NTD-RBD-Δ13)	106	107

Table 11B. Domain Fusion C-Terminal Truncation

SARS-CoV-2 NTD-RBD with C-terminal Truncation of 13 Amino Acids (NTD-RBD-Δ13)		
SEQ ID NO: 105 consists of from 5' end to 3' end: 5' UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 106 and 3' UTR SEQ ID NO: 4.		105
Chemistry	1-methylpseudouridine	
Cap	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCG GCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGUGAGCAGCCAGUGCG UGAACCUGACCACCCGGACCCAGCUGCCACCAGCCUACACCAACAG CUUCACCCGGGGCGUCUACUACCCCGACAAGGUGUUCCGGAGCAGC GUCCUGCACAGCACCCAGGACCUGUUCCUGCCCUUCUUCAGCAACG UGACCUGGUUCCACGCCAUCCACGUGAGCGGCACCAACGGCACCAA GCGGUUCGACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUC GCCAGCACCGAGAAGAGCAACAUCAUCCGGGGCUGGAUCUUCGGCA CCACCCUGGACAGCAAGACCCAGAGCCUGCUGAUCGUGAAUAACGC CACCAACGUGGUGAUCAAGGUGUGCGAGUUCCAGUUCUGCAACGAC CCCUUCCUGGGCGUGUACUACCACAAGAACAACAAGAGCUGGAUGG AGAGCGAGUUCCGGGUGUACAGCAGCGCCAACAACUGCACCUUCGA GUACGUGAGCCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAGGGC AACUUCAAGAACCUGCGGGAGUUCGUGUUCAAGAACAUCGACGGCU ACUUCAAGAUCUACAGCAAGCACACCCCAAUCAACCUGGUGCGGGA UCUGCCCCAGGGCUUCUCAGCCCUGGAGCCCCUGGUGGACCUGCCC AUCGGCAUCAACAUCACCCGGUUCCAGACCCUGCUGGCCCUGCACC GGAGCUACCUGACCCCAGGCGACAGCAGCAGCGGGUGGACAGCAGG CGCGGCUGCUUACUACGUGGGCUACCUGCAGCCCCGGACCUUCCUG CUGAAGUACAACGAGAACGGCACCAUCACCGACGCCGUGGACGGAG GCGGAUCGGGAGGCGGACCCAACAUCACCAACCUGUGCCCCUUCGG CGAGGUGUUCAACGCCACCCGGUUCGCCAGCGUGUACGCCUGGAAC CGGAAGCGGAUCAGCAACUGCGUGGCCGACUACAGCGUGCUGUACA ACAGCGCCAGCUUCAGCACCUUCAAGUGCUACGGCGUGAGCCCCAC CAAGCUGAACGACCUGUGCUUCACCAACGUGUACGCCGACAGCUUC	106

	GUGAUCCGUGGCGACGAGGUGCGGCAGAUCGCACCCGGCCAGACAG GCAAGAUCGCCGACUACAACUACAAGCUGCCCGACGACUUCACCGG CUGCGUGAUCGCCUGGAACAGCAACAACCUCGACAGCAAGGUGGGC GGCAACUACAACUACCUGUACCGGCUGUUCCGGAAGAGCAACCUGA AGCCCUUCGAGCGGGACAUCAGCACCGAGAUCUACCAAGCCGGCUC CACCCCUUGCAACGGCGUGGAGGGCUUCAACUGCUACUUCCCUCUG CAGAGCUACGGCUUCCAGCCCACCAACGGCGUGGGCUACCAGCCCU ACCGGGUGGUGGUGCUGAGCUUCGAGCUGCUGCACGCCCCAGCCAC CGUGUGUGGCCCCAAGUCUGGCGGAGGCAGCAUCCUGGCCAUCUAC AGCCUGGGCUUCAUCGCCGGCCUGAUCGCCAUCGUGAUGGUGACCA UCAUGCUGUGCUGCAUGACCAGCUGCUGCAGCUGCCUGAAGGGCUG UUGCAGCUGCGGCAGCUGCUGCAAGUUCGACGAGGACGAC	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGG CCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGC	4
Corresponding amino acid sequence	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSS VLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYF ASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCND PFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQG NFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLP IGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFL LKYNENGTITDAVDgggsgggPNITNLCPFGEVFNATRFASVYAWN RKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSF VIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVG GNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPL QSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKsgggsilaiy sLGFIAGLIAIVMVTIMLCCMTSCCSCLKGCCSCGSCCKFDEDD	107
PolyA tail	100 nt	

Domain Extensions

SARS-CoV-2 S protein domain antigens, in some embodiments, include "extended" regions that include sequences adjacent to and/or flanking what is understood in the art to be the NTD domain or the RBD domain. The RBD_EXT series encompasses the SD1 (subdomain 1). The NTD_EXT series encompasses a C-terminal helix in the NTD. Some B cells and antibodies recognize conformational epitopes found only in properly folded, but not denatured, forms of the SARS-CoV-2 S protein NTD and RBD. Inclusion of sequences adjacent to and/or flanking the NTD and RBD domains not only can provide additional B-cell epitopes to the antigen, but may potentially result in more optimal folding of those domains and stimulate B cells with antibodies specific to epitopes that may be found on the edge of either domain. Furthermore, the inclusion of these extension sequences may thus increase the distance between the NTD or RBD and the expressing cell membrane, increasing exposure of both domains to antibodies that may bind less efficiently if the expressed protein was too close to the cell surface. Finally, the inclusion of extension sequences increases the pool of peptides that could potentially be presented to $CD4^+$ T cells by B cells that have recognized an NTD or RBD epitope, then processed the entire protein for antigen presentation, thereby increasing the chance that an NTD or RBD-specific B cell receives sufficient T cell help. Non-limiting example of SARS-CoV-2 domain extensions and the mRNA encoding them are provided in **Tables 12A** and **12B** below.

Table 12. Domain Extensions

Name	SEQ ID NO:	
	mRNA ORF	Protein
SARS-CoV-2 NTD DS Extended Linked to Transmembrane Domain (NTD-EXT-F43C-TM)	52	53
SARS-CoV-2 NTD DS Extended Linked to Transmembrane Domain (NTD-F43C-EXT-TM)	55	56
SARS-CoV-2 NTD Extended Linked to Transmembrane Domain (NTD-EXT-TM)	58	59
SARS-CoV-2 RBD Extended Linked to Transmembrane Domain (RBD-EXT-TM)	85	86
SARS-CoV-2 RBD DS Extended Linked to Transmembrane Domain (RBD-Q563D-EXT-TM)	88	89
SARS-CoV-2 NTD-RBD Extended Linked to Transmembrane Domain (NTD-RBD-EXT-TM)	115	116
SARS-CoV-2 NTD Extended-RBD Linked to Transmembrane Domain (NTD-EXT-RBD-TM)	118	119
SARS-CoV-2 NTD Extended-RBD-Extended Linked to Transmembrane Domain (NTD-EXT-RBD-EXT-TM)	121	122

Table 12B. Domain Extensions

SARS-CoV-2 NTD DS Extended Linked to Transmembrane Domain (NTD-EXT-F43C-TM)		
SEQ ID NO: 51 consists of from 5' end to 3' end: 5' UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 52 and 3' UTR SEQ ID NO: 4.		51
Chemistry	1-methylpseudouridine	
Cap	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCG GCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGUGAGCAGCCAGUGCG UGAACCUGACCACCCGGACCCAGCUGCCACCAGCCUACACCAACAG CUUCACCCGGGGCGUCUACUACCCCGACAAGGUGUGCCGGAGCAGC GUCCUGCACAGCACCCAGGACCUGUUCCUGCCCUUCUUCAGCAACG UGACCUGGUUCCACGCCAUCCACGUGAGCGGCACCAACGGCACCAA GCGGUUCGACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUC GCCAGCACCGAGAAGAGCAACAUCAUCCGGGGCUGGAUCUUCGGCA CCACCCUGGACAGCAAGACCCAGAGCCUGCUGAUCGUGAAUAACGC CACCAACGUGGUGAUCAAGGUGUGCGAGUUCCAGUUCUGCAACGAC CCCUUCCUGGGCGUGUACUACCACAAGAACAACAAGAGCUGGAUGG AGAGCGAGUUCCGGGUGUACAGCAGCGCCAACAACUGCACCUUCGA GUACGUGAGCCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAGGGC AACUUCAAGAACCUGCGGGAGUUCGUGUUCAAGAACAUCGACGGCU ACUUCAAGAUCUACAGCAAGCACACCCCAAUCAACCUGGUGCGGGA UCUGCCCCAGGGCUUCUCAGCCCUGGAGCCCCUGGUGGACCUGCCC AUCGGCAUCAACAUCACCCGGUUCCAGACCCUGCUGGCCCUGCACC GGAGCUACCUGACCCCAGGCGACAGCAGCAGCGGGUGGACAGCAGG CGCGGCUGCUUACUACGUGGGCUACCUGCAGCCCCGGACCUUCCUG CUGAAGUACAACGAGAACGGCACCAUCACCGACGCCGUGGACUCUG GCGGAGGCAGCAUCCUGGCCAUCUACAGCACCGUGGCCAGCAGCCU GGUGCUGCUGGUGAGCCUGGGCGCCAUCAGCUUC	52
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGG CCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGC	4

Corresponding amino acid sequence	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVCRSS VLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYF ASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCND PFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQG NFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLP IGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFL LKYNENGTITDAVDsgggsilaiystvasslvllvslgaisf	53
PolyA tail	100 nt	
SARS-CoV-2 NTD DS Extended Linked to Transmembrane Domain (NTD-F43C-EXT-TM)		
SEQ ID NO: 54 consists of from 5' end to 3' end: 5' UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 55 and 3' UTR SEQ ID NO: 4.		54
Chemistry	1-methylpseudouridine	
Cap	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCG GCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGUGAGCAGCCAGUGCG UGAACCUGACCACCCGGACCCAGCUGCCACCAGCCUACACCAACAG CUUCACCCGGGGCGUCUACUACCCCGACAAGGUGUGCCGGAGCAGC GUCCUGCACAGCACCCAGGACCUGUUCCUGCCCUUCUUCAGCAACG UGACCUGGUUCCACGCCAUCCACGUGAGCGGCACCAACGGCACCAA GCGGUUCGACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUC GCCAGCACCGAGAAGAGCAACAUCAUCCGGGGCUGGAUCUUCGGCA CCACCCUGGACAGCAAGACCCAGAGCCUGCUGAUCGUGAAUAACGC CACCAACGUGGUGAUCAAGGUGUGCGAGUUCCAGUUCUGCAACGAC CCCUUCCUGGGCGUGUACUACCACAAGAACAACAAGAGCUGGAUGG AGAGCGAGUUCCGGGUGUACAGCAGCGCCAACAACUGCACCUUCGA GUACGUGAGCCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAGGGC AACUUCAAGAACCUGCGGGAGUUCGUGUUCAAGAACAUCGACGGCU ACUUCAAGAUCUACAGCAAGCACACCCCAAUCAACCUGGUGCGGGA UCUGCCCCAGGGCUUCUCAGCCCUGGAGCCCCUGGUGGACCUGCCC AUCGGCAUCAACAUCACCCGGUUCCAGACCCUGCUGGCCCUGCACC GGAGCUACCUGACCCCAGGCGACAGCAGCAGCGGGUGGACAGCAGG CGCGGCUGCUUACUACGUGGGCUACCUGCAGCCCCGGACCUUCCUG CUGAAGUACAACGAGAACGGCACCAUCACCGACGCCGUGGACUGCG CCCUGGACCCUCUGAGCGAGACCAAGUGCACCCUGAAGAGCUUCAC CUCUGGCGGAGGCAGCAUCCUGGCCAUCUACAGCACCGUGGCCAGC AGCCUGGUGCUGCUGGUGAGCCUGGGCGCCAUCAGCUUC	55
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGG CCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGC	4
Corresponding amino acid sequence	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVCRSS VLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYF ASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCND PFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQG NFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLP IGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFL LKYNENGTITDAVDCALDPLSETKCTLKSFTsgggsilaiystvas slvllvslgaisf	56
PolyA tail	100 nt	
SARS-CoV-2 NTD Extended Linked to Transmembrane Domain (NTD-EXT-TM)		
SEQ ID NO: 57 consists of from 5' end to 3' end: 5' UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 58 and 3' UTR SEQ ID NO: 4.		57
Chemistry	1-methylpseudouridine	
Cap	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCG GCGCCGCCACC	2

ORF of mRNA Construct (excluding the stop codon)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGUGAGCAGCCAGUGCG UGAACCUGACCACCCGGACCCAGCUGCCACCAGCCUACACCAACAG CUUCACCCGGGGCGUCUACUACCCCGACAAGGUGUUCCGGAGCAGC GUCCUGCACAGCACCCAGGACCUGUUCCUGCCCUUCUUCAGCAACG UGACCUGGUUCCACGCCAUCCACGUGAGCGGCACCAACGGCACCAA GCGGUUCGACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUC GCCAGCACCGAGAAGAGCAACAUCAUCCGGGGCUGGAUCUUCGGCA CCACCCUGGACAGCAAGACCCAGAGCCUGCUGAUCGUGAAUAACGC CACCAACGUGGUGAUCAAGGUGUGCGAGUUCCAGUUCUGCAACGAC CCCUUCCUGGGCGUGUACUACCACAAGAACAACAAGAGCUGGAUGG AGAGCGAGUUCCGGGUGUACAGCAGCGCCAACAACUGCACCUUCGA GUACGUGAGCCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAGGGC AACUUCAAGAACCUGCGGGAGUUCGUGUUCAAGAACAUCGACGGCU ACUUCAAGAUCUACAGCAAGCACACCCCAAUCAACCUGGUGCGGGA UCUGCCCCAGGGCUUCUCAGCCCUGGAGCCCCUGGUGGACCUGCCC AUCGGCAUCAACAUCACCCGGUUCCAGACCCUGCUGGCCCUGCACC GGAGCUACCUGACCCCAGGCGACAGCAGCAGCGGGUGGACAGCAGG CGCGGCUGCUUACUACGUGGGCUACCUGCAGCCCCGGACCUUCCUG CUGAAGUACAACGAGAACGGCACCAUCACCGACGCCGUGGACUGCG CCCUGGACCCUCUGAGCGAGACCAAGUGCACCCUGAAGAGCUUCAC CUCUGGCGGAGGCAGCAUCCUGGCCAUCUACAGCACCGUGGCCAGC AGCCUGGUGCUGCUGGUGAGCCUGGGCGCCAUCAGCUUC	58
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGG CCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGC	4
Corresponding amino acid sequence	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSS VLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYF ASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCND PFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQG NFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLP IGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFL LKYNENGTITDAVDCALDPLSETKCTLKSFTsgggsilaiystvas slvllvslgaisf	59
PolyA tail	100 nt	
SARS-CoV-2 RBD Extended Linked to Transmembrane Domain (RBD-EXT-TM)		
SEQ ID NO: 84 consists of from 5' end to 3' end: 5' UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 85 and 3' UTR SEQ ID NO: 4.		84
Chemistry	1-methylpseudouridine	
Cap	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCG GCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUACAGCAUGCAGCUGGCUAGCUGCGUGACCCUGACCCUGGUGC UGCUGGUGAACAGCCAGCGGGUGCAGCCCACCGAGAGCAUCGUGCG GUUCCCCAACAUCACCAACCUGUGCCCCUUCGGCGAGGUGUUCAAC GCCACCCGGUUCGCCAGCGUGUACGCCUGGAACCGGAAGCGGAUCA GCAACUGCGUGGCCGACUACAGCGUGCUGUACAACAGCGCCAGCUU CAGCACCUUCAAGUGCUACGGCGUGAGCCCCACCAAGCUGAACGAC CUGUGCUUCACCAACGUGUACGCCGACAGCUUCGUGAUCCGUGGCG ACGAGGUGCGGCAGAUCGCACCCGGCCAGACAGGCAAGAUCGCCGA CUACAACUACAAGCUGCCCGACGACUUCACCGGCUGCGUGAUCGCC UGGAACAGCAACAACCUCGACAGCAAGGUGGGCGGCAACUACAACU ACCUGUACCGGCUGUUCCGGAAGAGCAACCUGAAGCCCUUCGAGCG GGACAUCAGCACCGAGAUCUACCAAGCCGGCUCCACCCCUUGCAAC GGCGUGGAGGGCUUCAACUGCUACUUCCCUCUGCAGAGCUACGGCU UCCAGCCCACCAACGGCGUGGGCUACCAGCCCUACCGGGUGGUGGU GCUGAGCUUCGAGCUGCUGCACGCCCCAGCCACCGUGUGUGGCCCC AAGAAGAGCACCAACCUGGUGAAGAACAAGUGCGUGAACUUCAACU UCAACGGCCUUACCGGCACCGGCGUGCUGACCGAGAGCAACAAGAA AUUCCUGCCCUUUCAGCAGUUCGGCCGGGACAUCGCCGACACCACC GACGCUGUGCGGGAUCCCCAGACCCUGGAGAUCCUGGACAUCACCC	85

	CUUGCAGCUCUGGCGGAGGCAGCAUCCUGGCCAUCUACAGCACCGU GGCCAGCAGCCUGGUGCUGCUGGUGAGCCUGGGCGCCAUCAGCUUC	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGG CCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGC	4
Corresponding amino acid sequence	mysmqlascvtltlvllvnsQRVQPTESIVRFPNITNLCPFGEVFN ATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLND LCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIA WNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCN GVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGP KKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTT DAVRDPQTLEILDITPCSsgggsilaiystvasslvllvslgaisf	86
PolyA tail	100 nt	
SARS-CoV-2 RBD DS Extended Linked to Transmembrane Domain (RBD-Q563D-EXT-TM)		
SEQ ID NO: 87 consists of from 5' end to 3' end: 5' UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 88 and 3' UTR SEQ ID NO: 4.		87
Chemistry	1-methylpseudouridine	
Cap	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCG GCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUACAGCAUGCAGCUGGCUAGCUGCGUGACCCUGACCCUGGUGC UGCUGGUGAACAGCCAGCGGGUGCAGCCCACCGAGAGCAUCGUGCG GUUCCCCAACAUCACCAACCUGUGCCCCUUCGGCGAGGUGUUCAAC GCCACCCGGUUCGCCAGCGUGUACGCCUGGAACCGGAAGCGGAUCA GCAACUGCGUGGCCGACUACAGCGUGCUGUACAACAGCGCCAGCUU CAGCACCUUCAAGUGCUACGGCGUGAGCCCCACCAAGCUGAACGAC CUGUGCUUCACCAACGUGUACGCCGACAGCUUCGUGAUCCGUGGCG ACGAGGUGCGGCAGAUCGCACCCGGCCAGACAGGCAAGAUCGCCGA CUACAACUACAAGCUGCCCGACGACUUCACCGGCUGCGUGAUCGCC UGGAACAGCAACAACCUCGACAGCAAGGUGGGCGGCAACUACAACU ACCUGUACCGGCUGUUCCGGAAGAGCAACCUGAAGCCCUUCGAGCG GGACAUCAGCACCGAGAUCUACCAAGCCGGCUCCACCCCUUGCAAC GGCGUGGAGGGCUUCAACUGCUACUUCCCCUCUGCAGAGCUACGGCU UCCAGCCCACCAACGGCGUGGGCUACCAGCCCUACCGGGUGGUGGU GCUGAGCUUCGAGCUGCUGCACGCCCCAGCCACCGUGUGUGGCCCC AAGAAGAGCACCAACCUGGUGAAGAACAAGUGCGUGAACUUCAACU UCAACGGCCUUACCGGCACCGGCGUGCUGACCGAGAGCAACAAGAA AUUCCUGCCCUUUUGCCAGUUCGGCCGGGACAUCGCCGACACCACC GACGCUGUGCGGGAUCCCCAGACCCUGGAGAUCCUGGACAUCACCC CUUGCAGCUCUGGCGGAGGCAGCAUCCUGGCCAUCUACAGCACCGU GGCCAGCAGCCUGGUGCUGCUGGUGAGCCUGGGCGCCAUCAGCUUC	88
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGG CCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGC	4
Corresponding amino acid sequence	mysmqlascvtltlvllvnsQRVQPTESIVRFPNITNLCPFGEVFN ATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLND LCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIA WNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCN GVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGP KKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFCQFGRDIADTT DAVRDPQTLEILDITPCSsgggsilaiystvasslvllvslgaisf	89
PolyA tail	100 nt	
SARS-CoV-2 NTD-RBD Extended Linked to Transmembrane Domain (NTD-RBD-EXT-TM)		
SEQ ID NO: 114 consists of from 5' end to 3' end: 5' UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 115 and 3' UTR SEQ ID NO: 4.		114
Chemistry	1-methylpseudouridine	
Cap	7mG(5')ppp(5')NlmpNp	

5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCG GCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGUGAGCAGCCAGUGCG UGAACCUGACCACCCGGACCCAGCUGCCACCAGCCUACACCAACAG CUUCACCCGGGGCGUCUACUACCCCGACAAGGUGUUCCGGAGCAGC GUCCUGCACAGCACCCAGGACCUGUUCCUGCCCUUCUUCAGCAACG UGACCUGGUUCCACGCCAUCCACGUGAGCGGCACCAACGGCACCAA GCGGUUCGACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUC GCCAGCACCGAGAAGAGCAACAUCAUCCGGGGCUGGAUCUUCGGCA CCACCCUGGACAGCAAGACCCAGAGCCUGCUGAUCGUGAAUAACGC CACCAACGUGGUGAUCAAGGUGUGCGAGUUCCAGUUCUGCAACGAC CCCUUCCUGGGCGUGUACUACCACAAGAACAACAAGAGCUGGAUGG AGAGCGAGUUCCGGGUGUACAGCAGCGCCAACAACUGCACCUUCGA GUACGUGAGCCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAGGGC AACUUCAAGAACCUGCGGGAGUUCGUGUUCAAGAACAUCGACGGCU ACUUCAAGAUCUACAGCAAGCACACCCCAAUCAACCUGGUGCGGGA UCUGCCCCAGGGCUUCUCAGCCCUGGAGCCCCUGGUGGACCUGCCC AUCGGCAUCAACAUCACCCGGUUCCAGACCCUGCUGGCCCUGCACC GGAGCUACCUGACCCCAGGCGACAGCAGCAGCGGGUGGACAGCAGG CGCGGCUGCUUACUACGUGGGCUACCUGCAGCCCCGGACCUUCCUG CUGAAGUACAACGAGAACGGCACCAUCACCGACGCCGUGGACGGAG GCGGAUCGGGAGGCGGACAGCGGGUGCAGCCCACCGAGAGCAUCGU GCGGUUCCCCAACAUCACCAACCUGUGCCCCUUCGGCGAGGUGUUC AACGCCACCCGGUUCGCCAGCGUGUACGCCUGGAACCGGAAGCGGA UCAGCAACUGCGUGGCCGACUACAGCGUGCUGUACAACAGCGCCAG CUUCAGCACCUUCAAGUGCUACGGCGUGAGCCCCACCAAGCUGAAC GACCUGUGCUUCACCAACGUGUACGCCGACAGCUUCGUGAUCCGUG GCGACGAGGUGCGGCAGAUCGCACCCGGCCAGACAGGCAAGAUCGC CGACUACAACUACAAGCUGCCCGACGACUUCACCGGCUGCGUGAUC GCCUGGAACAGCAACAACCUCGACAGCAAGGUGGGCGGCAACUACA ACUACCUGUACCGGCUGUUCCGGAAGAGCAACCUGAAGCCCUUCGA GCGGGACAUCAGCACCGAGAUCUACCAAGCCGGCUCCACCCCUUGC AACGGCGUGGAGGGCUUCAACUGCUACUUCCCUCUGCAGAGCUACG GCUUCCAGCCCACCAACGGCGUGGGCUACCAGCCCUACCGGGUGGU GGUGCUGAGCUUCGAGCUGCUGCACGCCCCAGCCACCGUGUGUGGC CCCAAGAAGAGCACCAACCUGGUGAAGAACAAGUGCGUGAACUUCA ACUUCAACGGCCUUACCGGCACCGGCGUGCUGACCGAGAGCAACAA GAAAUUCCUGCCCUUUCAGCAGUUCGGCCGGGACAUCGCCGACACC ACCGACGCUGUGCGGGAUCCCCAGACCCUGGAGAUCCUGGACAUCA CCCCUUGCAGCUCUGGCGGAGGCAGCAUCCUGGCCAUCUACAGCAC CGUGGCCAGCAGCCUGGUGCUGCUGGUGAGCCUGGGCGCCAUCAGC UUC	115
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGG CCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGC	4
Corresponding amino acid sequence	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSS VLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYF ASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCND PFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQG NFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLP IGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFL LKYNENGTITDAVDgggsgggQRVQPTESIVRFPNITNLCPFGEVF NATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLN DLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVI AWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPC NGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCG PKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADT TDAVRDPQTLEILDITPCSsgggsilaiystvasslvllvslgais f	116
PolyA tail	100 nt	
SARS-CoV-2 NTD Extended-RBD Linked to Transmembrane Domain (NTD-EXT-RBD-TM)		

SEQ ID NO: 117 consists of from 5' end to 3' end: 5' UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 118 and 3' UTR SEQ ID NO: 4.		117
Chemistry	1-methylpseudouridine	
Cap	7mG(5')ppp(5')NImpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCG GCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGUGAGCAGCCAGUGCG UGAACCUGACCACCCGGACCCAGCUGCCACCAGCCUACACCAACAG CUUCACCCGGGGCGUCUACUACCCCGACAAGGUGUUCCGGAGCAGC GUCCUGCACAGCACCCAGGACCUGUUCCUGCCCUUCUUCAGCAACG UGACCUGGUUCCACGCCAUCCACGUGAGCGGCACCAACGGCACCAA GCGGUUCGACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUC GCCAGCACCGAGAAGAGCAACAUCAUCCGGGGCUGGAUCUUCGGCA CCACCCUGGACAGCAAGACCCAGAGCCUGCUGAUCGUGAAUAACGC CACCAACGUGGUGAUCAAGGUGUGCGAGUUCCAGUUCUGCAACGAC CCCUUCCUGGGCGUGUACUACCACAAGAACAACAAGAGCUGGAUGG AGAGCGAGUUCCGGGUGUACAGCAGCGCCAACAACUGCACCUUCGA GUACGUGAGCCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAGGGC AACUUCAAGAACCUGCGGGAGUUCGUGUUCAAGAACAUCGACGGCU ACUUCAAGAUCUACAGCAAGCACACCCCAAUCAACCUGGUGCGGGA UCUGCCCCAGGGCUUCUCAGCCCUGGAGCCCCUGGUGGACCUGCCC AUCGGCAUCAACAUCACCCGGUUCCAGACCCUGCUGGCCCUGCACC GGAGCUACCUGACCCCAGGCGACAGCAGCAGCGGGUGGACAGCAGG CGCGGCUGCUUACUACGUGGGCUACCUGCAGCCCCGGACCUUCCUG CUGAAGUACAACGAGAACGGCACCAUCACCGACGCCGUGGACUGCG CCCUGGACCCUCUGAGCGAGACCAAGUGCACCCUGAAGAGCUUCAC CGGAGGCGGAUCGGGAGGCGGACCCAACAUCACCAACCUGUGCCCC UUCGGCGAGGUGUUCAACGCCACCCGGUUCGCCAGCGUGUACGCCU GGAACCGGAAGCGGAUCAGCAACUGCGUGGCCGACUACAGCGUGCU GUACAACAGCGCCAGCUUCAGCACCUUCAAGUGCUACGGCGUGAGC CCCACCAAGCUGAACGACCUGUGCUUCACCAACGUGUACGCCGACA GCUUCGUGAUCCGUGGCGACGAGGUGCGGCAGAUCGCACCCGGCCA GACAGGCAAGAUCGCCGACUACAACUACAAGCUGCCCGACGACUUC ACCGGCUGCGUGAUCGCCUGGAACAGCAACAACCUCGACAGCAAGG UGGGCGGCAACUACAACUACCUGUACCGGCUGUUCCGGAAGAGCAA CCUGAAGCCCUUCGAGCGGGACAUCAGCACCGAGAUCUACCAAGCC GGCUCCACCCCUUGCAACGGCGUGGAGGGCUUCAACUGCUACUUCC CUCUGCAGAGCUACGGCUUCCAGCCCACCAACGGCGUGGGCUACCA GCCCUACCGGGUGGUGGUGCUGAGCUUCGAGCUGCUGCACGCCCCA GCCACCGUGUGUGGCCCCAAGUCUGGCGGAGGCAGCAUCCUGGCCA UCUACAGCACCGUGGCCAGCAGCCUGGUGCUGCUGGUGAGCCUGGG CGCCAUCAGCUUC	118
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGG CCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGC	4
Corresponding amino acid sequence	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSS VLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYF ASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCND PFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQG NFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLP IGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFL LKYNENGTITDAVDCALDPLSETKCTLKSFTgggsgggPNITNLCP FGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVS PTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDF TGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQA GSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFELLHAP ATVCGPKsgggsilaiystvasslvllvslgaisf	119
PolyA tail	100 nt	
SARS-CoV-2 NTD Extended-RBD-Extended Linked to Transmembrane Domain (NTD-EXT-RBD-EXT-TM)		

SEQ ID NO: 120 consists of from 5' end to 3' end: 5' UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 121 and 3' UTR SEQ ID NO: 4.		120
Chemistry	1-methylpseudouridine	
Cap	7mG(5')ppp(5')NImpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCG GCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGUGAGCAGCCAGUGCG UGAACCUGACCACCCGGACCCAGCUGCCACCAGCCUACACCAACAG CUUCACCCGGGGCGUCUACUACCCCGACAAGGUGUUCCGGAGCAGC GUCCUGCACAGCACCCAGGACCUGUUCCUGCCCUUCUUCAGCAACG UGACCUGGUUCCACGCCAUCCACGUGAGCGGCACCAACGGCACCAA GCGGUUCGACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUC GCCAGCACCGAGAAGAGCAACAUCAUCCGGGGCUGGAUCUUCGGCA CCACCCUGGACAGCAAGACCCAGAGCCUGCUGAUCGUGAAUAACGC CACCAACGUGGUGAUCAAGGUGUGCGAGUUCCAGUUCUGCAACGAC CCCUUCCUGGGCGUGUACUACCACAAGAACAACAAGAGCUGGAUGG AGAGCGAGUUCCGGGUGUACAGCAGCGCCAACAACUGCACCUUCGA GUACGUGAGCCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAGGGC AACUUCAAGAACCUGCGGGAGUUCGUGUUCAAGAACAUCGACGGCU ACUUCAAGAUCUACAGCAAGCACACCCCAAUCAACCUGGUGCGGGA UCUGCCCCAGGGCUUCUCAGCCCUGGAGCCCCUGGUGGACCUGCCC AUCGGCAUCAACAUCACCCGGUUCCAGACCCUGCUGGCCCUGCACC GGAGCUACCUGACCCCAGGCGACAGCAGCAGCGGGUGGACAGCAGG CGCGGCUGCUUACUACGUGGGCUACCUGCAGCCCCGGACCUUCCUG CUGAAGUACAACGAGAACGGCACCAUCACCGACGCCGUGGACUGCG CCCUGGACCCUCUGAGCGAGACCAAGUGCACCCUGAAGAGCUUCAC CGGAGGCGGAUCGGGAGGCGGACAGCGGGUGCAGCCCACCGAGAGC AUCGUGCGGUUCCCCAACAUCACCAACCUGUGCCCCUUCGGCGAGG UGUUCAACGCCACCCGGUUCGCCAGCGUGUACGCCUGGAACCGGAA GCGGAUCAGCAACUGCGUGGCCGACUACAGCGUGCUGUACAACAGC GCCAGCUUCAGCACCUUCAAGUGCUACGGCGUGAGCCCCACCAAGC UGAACGACCUGUGCUUCACCAACGUGUACGCCGACAGCUUCGUGAU CCGUGGCGACGAGGUGCGGCAGAUCGCACCCGGCCAGACAGGCAAG AUCGCCGACUACAACUACAAGCUGCCCGACGACUUCACCGGCUGCG UGAUCGCCUGGAACAGCAACAACCUCGACAGCAAGGUGGGCGGCAA CUACAACUACCUGUACCGGCUGUUCCGGAAGAGCAACCUGAAGCCC UUCGAGCGGGACAUCAGCACCGAGAUCUACCAAGCCGGCUCCACCC CUUGCAACGGCGUGGAGGGCUUCAACUGCUACUUCCCUCUGCAGAG CUACGGCUUCCAGCCCACCAACGGCGUGGGCUACCAGCCCUACCGG GUGGUGGUGCUGAGCUUCGAGCUGCUGCACGCCCCAGCCACCGUGU GUGGCCCCAAGAAGAGCACCAACCUGGUGAAGAACAAGUGCGUGAA CUUCAACUUCAACGGCCUUACCGGCACCGGCGUGCUGACCGAGAGC AACAAGAAAUUCCUGCCCUUUCAGCAGUUCGGCCGGGACAUCGCCG ACACCACCGACGCUGUGCGGGAUCCCCAGACCCUGGAGAUCCUGGA CAUCACCCCUUGCAGCUCUGGCGGAGGCAGCAUCCUGGCCAUCUAC AGCACCGUGGCCAGCAGCCUGGUGCUGCUGGUGAGCCUGGGCGCCA UCAGCUUC	121
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGG CCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGC	4
Corresponding amino acid sequence	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSS VLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYF ASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCND PFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQG NFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLP IGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFL LKYNENGTITDAVDCALDPLSETKCTLKSFTgggsgggQRVQPTES IVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNS ASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGK IADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKP FERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYR	122

	VVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTES NKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSsgggsilaiy stvasslvllvslgaisf	
PolyA tail	100 nt	

Domain Mixtures

The present disclosure provides, in some aspects, compositions that comprise a mixture of mRNAs encoding SARS-CoV-2 S protein subdomains. In one example, a composition comprises a mixture of an mRNA encoding an NTD (with or without SD1, SD2, and/or a transmembrane domain) and an mRNA encoding an RBD (with or without SD1, SD2, and/or a transmembrane domain). In some embodiments, a composition comprises an mRNA (e.g., SEQ ID NO: 45 or 46) encoding an NTD linked to a transmembrane domain (e.g., SEQ ID NO: 47) and an mRNA (e.g., SEQ ID NO: 75 or 76 encoding an RBD linked to a transmembrane domain (e.g., SEQ ID NO: 77).

The ratio of the concentration of one mRNA to another in a composition may be 1:1 (50:50), 1:2, 1:3, 1:4, or 1:5. In some embodiments, the ratio is 1:1. For example, a composition may comprise a 1:1 ratio of an mRNA (e.g., SEQ ID NO: 45 or 46) encoding an NTD linked to a transmembrane domain (e.g., SEQ ID NO: 47) to an mRNA (e.g., SEQ ID NO: 75 or 76 encoding an RBD linked to a transmembrane domain (e.g., SEQ ID NO: 77). In some embodiments, the ratio is 1:2. For example, a composition may comprise a 1:2 ratio of an mRNA (e.g., SEQ ID NO: 45 or 46) encoding an NTD linked to a transmembrane domain (e.g., SEQ ID NO: 47) to an mRNA (e.g., SEQ ID NO: 75 or 76) encoding an RBD linked to a transmembrane domain (e.g., SEQ ID NO: 77). Another example, a composition may comprise a 1:2 ratio of an mRNA (e.g., SEQ ID NO: 75 or 76) encoding an RBD linked to a transmembrane domain (e.g., SEQ ID NO: 77) to an mRNA (e.g., SEQ ID NO: 45 or 46) encoding an NTD linked to a transmembrane domain (e.g., SEQ ID NO: 47). Different mRNAs encoding different antigens may stimulate immune responses of varying strength (Magini D et al. PLoS ONE. 2016; 11:e0161193), and administration of an equimolar ratio of two mRNAs encoding two different antigens may result in an immune response to one but not the other (John S et al. Vaccine. 2018; 36:1689 – 1699). Manipulation of the ratio of co-delivered mRNAs may be useful for eliciting broad immune responses that target desired antigens with equal potency.

Encoded Nanoparticle Antigens

The mRNA vaccines provided herein, in some embodiments, encode fusion proteins that comprise coronavirus antigens linked to a scaffold domain. In some embodiments, a scaffold domain imparts desired properties to an antigen encoded by an mRNA of the disclosure. For

example, scaffold domain may improve the immunogenicity of an antigen, e.g., by altering the structure of the antigen, altering the uptake and processing of the antigen, and/or causing the antigen to bind to another molecule. In some embodiments, a scaffold domain linked to antigen facilitates self-assembly of the antigen into a viral nanoparticle or a larger protein-folded immunogen. Non-limiting examples of scaffold domains that may be used as provide herein include, ferritin domains, lumazine synthetase domains, foldon domains, and encapsulin domains. Other scaffold domains may be used.

Ferritin

In some embodiments, a ferritin domain is used as a scaffold domain. Ferritin is a protein, the main function of which is intracellular iron storage. Ferritin is comprised of twenty-four (24) subunits, each composed of a four-alpha-helix bundle that self-assemble into a quaternary structure with octahedral symmetry (Cho K. J. et al. J Mol Biol. 2009; 390: 83–98; (Granier T. et al. J Biol Inorg Chem. 2003; 8: 105–111; and Lawson D.M. et al. Nature. 1991; 349: 541–544). Ferritin self-assembles into nanoparticles with robust thermal and chemical stability. Enclosing antigens within ferritin nanoparticles in this manner is expected to both delay degradation of the antigen and aggregate individual antigens, with each nanoparticle containing twenty-four (24) antigen subunits. Aggregation of multiple copies of the same antigen enhances both antigen uptake and migration by dendritic cells, as well as more robust $CD4^+$ and $CD8^+$ T cell responses (Kastenmüller K et al. J Clin Invest. 2011; 121(5):1782-96). Thus, the ferritin nanoparticle is a well-suited platform for antigen presentation and vaccine development.

An mRNA provided herein, in some embodiments, encodes an RBD linked to a ferritin domain, for example, through a glycine (e.g., GGG) linker domain. Other linkers may be used.

In other embodiments, an mRNA provided herein encodes an S1 domain of an S protein linked to a ferritin domain, for example, through a glycine (e.g., GGG) linker. As indicated elsewhere herein, other linkers may be used.

Non-limiting examples of SARS-CoV-2 antigens linked to a ferritin domain and the mRNA encoding them are provided in **Tables 13A and 13B** below.

Table 13A. Antigens Linked to a Ferritin Domain

Name	SEQ ID NO:	
	mRNA ORF	Protein
SARS-CoV-2 S1 Subunit Linked to Ferritin (S1-Ferritin)	7	8
SARS-CoV-2 RBD Linked to Ferritin (RBD-Ferritin)	64	65

Table 13B. Antigens Linked to a Ferritin Domain

SARS-CoV-2 S1 Subunit Linked to Ferritin (S1-Ferritin)		
SEQ ID NO: 6 consists of from 5' end to 3' end: 5' UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 7 and 3' UTR SEQ ID NO: 4.		6
Chemistry	1-methylpseudouridine	
Cap	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCG GCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGUGAGCAGCCAGUGCG UGAACCUGACCACCCGGACCCAGCUGCCACCAGCCUACACCAACAG CUUCACCCGGGGCGUCUACUACCCCGACAAGGUGUUCCGGAGCAGC GUCCUGCACAGCACCCAGGACCUGUUCCUGCCCUUCUUCAGCAACG UGACCUGGUUCCACGCCAUCCACGUGAGCGGCACCAACGGCACCAA GCGGUUCGACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUC GCCAGCACCGAGAAGAGCAACAUCAUCCGGGGCUGGAUCUUCGGCA CCACCCUGGACAGCAAGACCCAGAGCCUGCUGAUCGUGAAUAACGC CACCAACGUGGUGAUCAAGGUGUGCGAGUUCCAGUUCUGCAACGAC CCCUUCCUGGGCGUGUACUACCACAAGAACAACAAGAGCUGGAUGG AGAGCGAGUUCCGGGUGUACAGCAGCGCCAACAACUGCACCUUCGA GUACGUGAGCCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAGGGC AACUUCAAGAACCUGCGGGAGUUCGUGUUCAAGAACAUCGACGGCU ACUUCAAGAUCUACAGCAAGCACACCCCAAUCAACCUGGUGCGGGA UCUGCCCCAGGGCUUCUCAGCCCUGGAGCCCCUGGUGGACCUGCCC AUCGGCAUCAACAUCACCCGGUUCCAGACCCUGCUGGCCCUGCACC GGAGCUACCUGACCCCAGGCGACAGCAGCAGCGGGUGGACAGCAGG CGCGGCUGCUUACUACGUGGGCUACCUGCAGCCCCGGACCUUCCUG CUGAAGUACAACGAGAACGGCACCAUCACCGACGCCGUGGACUGCG CCCUGGACCCUCUGAGCGAGACCAAGUGCACCCUGAAGAGCUUCAC CGUGGAGAAGGGCAUCUACCAGACCAGCAACUUCCGGGUGCAGCCC ACCGAGAGCAUCGUGCGGUUCCCCAACAUCACCAACCUGUGCCCCU UCGGCGAGGUGUUCAACGCCACCCGGUUCGCCAGCGUGUACGCCUG GAACCGGAAGCGGAUCAGCAACUGCGUGGCCGACUACAGCGUGCUG UACAACAGCGCCAGCUUCAGCACCUUCAAGUGCUACGGCGUGAGCC CCACCAAGCUGAACGACCUGUGCUUCACCAACGUGUACGCCGACAG CUUCGUGAUCCGUGGCGACGAGGUGCGGCAGAUCGCACCCGGCCAG ACAGGCAAGAUCGCCGACUACAACUACAAGCUGCCCGACGACUUCA CCGGCUGCGUGAUCGCCUGGAACAGCAACAACCUCGACAGCAAGGU GGGCGGCAACUACAACUACCUGUACCGGCUGUUCCGGAAGAGCAAC CUGAAGCCCUUCGAGCGGGACAUCAGCACCGAGAUCUACCAAGCCG GCUCCACCCCUUGCAACGGCGUGGAGGGCUUCAACUGCUACUUCCC UCUGCAGAGCUACGGCUUCCAGCCCACCAACGGCGUGGGCUACCAG CCCUACCGGGUGGUGGUGCUGAGCUUCGAGCUGCUGCACGCCCCAG CCACCGUGUGUGGCCCCAAGAAGAGCACCAACCUGGUGAAGAACAA GUGCGUGAACUUCAACUUCAACGGCCUUACCGGCACCGGCGUGCUG ACCGAGAGCAACAAGAAAUUCCUGCCCUUUCAGCAGUUCGGCCGGG ACAUCGCCGACACCACCGACGCUGUGCGGGAUCCCCAGACCCUGGA GAUCCUGGACAUCACCCCUUGCAGCUUCGGCGGCGUGAGCGUGAUC ACCCCAGGCACCAACACCAGCAACCAGGUGGCCGUGCUGUACCAGG ACGUGAACUGCACCGAGGUGCCCGUGGCCAUCCACGCCGACCAGCU GACACCCACCUGGCGGGUCUACAGCACCGGCAGCAACGUGUUCCAG ACCCGGGCCGGUUGCCUGAUCGGCGCCGAGCACGUGAACAACAGCU ACGAGUGCGACAUCCCCAUCGGCGCCGGCAUCUGUGCCAGCUACCA GACCCAGACCAAUUCAGGAGGAGGCAGCGGCGGCGAUAUCAUCAAG CUUCUGAACGAGCAAGUUAACAAGGAAAUGCAGAGCAGUAAUCUCU ACAUGAGCAUGAGCAGCUGGUGCUACACCCACUCCCUGGACGGAGC AGGCCUCUUCCUGUUCGACCACGCAGCCGAGGAGUACGAGCACGCU AAGAAGUUGAUCAUUUUCUUGAACGAGAACAACGUGCCCGUGCAGC UAACGUCAAUCAGCGCACCUGAGCACAAGUUCGAGGGCCUGACCCA GAUCUUCCAGAAGGCCUACGAACACGAACAGCACAUCUCCGAGAGC AUCAACAAUAUUGUGGAUCACGCUAUCAAGUCCAAGGACCACGCUA	7

	CCUUCAACUUCCUGCAGUGGUACGUGGCCGAGCAACAUGAGGAGGA GGUGCUGUUCAAGGACAUCCUGGACAAGAUCGAGCUGAUCGGUAAU GAGAAUCACGGCCUGUACCUGGCCGACCAGUACGUGAAGGGCAUCG CCAAGAGCCGGAAGUCAGGCUCA	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGG CCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGC	4
Corresponding amino acid sequence	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSS VLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYF ASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCND PFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQG NFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLP IGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFL LKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQP TESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVL YNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQ TGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSN LKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQ PYRVVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVL TESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVI TPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQ TRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSgggSGGDIIK LLNEQVNKEMQSSNLYMSMSSWCYTHSLDGAGLFLFDHAAEEYEHA KKLIIFLNENNVPVQLTSISAPEHKFEGLTQIFQKAYEHEQHISES INNIVDHAIKSKDHATFNFLQWYVAEQHEEEVLFKDILDKIELIGN ENHGLYLADQYVKGIAKSRKSGS	8
PolyA tail	100 nt	
SARS-CoV-2 RBD Linked to Ferritin (RBD-Ferritin)		
SEQ ID NO: 63 consists of from 5' end to 3' end: 5' UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 64 and 3' UTR SEQ ID NO: 4.		63
Chemistry	1-methylpseudouridine	
Cap	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCG GCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUACAGCAUGCAGCUGGCUAGCUGCGUGACCCUGACCCUGGUGC UGCUGGUGAACAGCCAGCCCAACAUCACCAACCUGUGCCCCUUCGG CGAGGUGUUCAACGCCACCCGGUUCGCCAGCGUGUACGCCUGGAAC CGGAAGCGGAUCAGCAACUGCGUGGCCGACUACAGCGUGCUGUACA ACAGCGCCAGCUUCAGCACCUUCAAGUGCUACGGCGUGAGCCCCAC CAAGCUGAACGACCUGUGCUUCACCAACGUGUACGCCGACAGCUUC GUGAUCCGUGGCGACGAGGUGCGGCAGAUCGCACCCGGCCAGACAG GCAAGAUCGCCGACUACAACUACAAGCUGCCCGACGACUUCACCGG CUGCGUGAUCGCCUGGAACAGCAACAACCUCGACAGCAAGGUGGGC GGCAACUACAACUACCUGUACCGGCUGUUCCGGAAGAGCAACCUGA AGCCCUUCGAGCGGGACAUCAGCACCGAGAUCUACCAAGCCGGCUC CACCCCUUGCAACGGCGUGGAGGGCUUCAACUGCUACUUCCCUCUG CAGAGCUACGGCUUCCAGCCCACCAACGGCGUGGGCUACCAGCCCU ACCGGGUGGUGGUGCUGAGCUUCGAGCUGCUGCACGCCCCAGCCAC CGUGUGUGGCCCCAAGGGAGGAGGCAGCGGCGGCGAUAUCAUCAAG CUUCUGAACGAGCAAGUUAACAAGGAAAUGCAGAGCAGUAAUCUCU ACAUGAGCAUGAGCAGCUGGUGCUACACCCACUCCCUGGACGGAGC AGGCCUCUUCCUGUUCGACCACGCAGCCGAGGAGUACGAGCACGCU AAGAAGUUGAUCAUUUUCUUGAACGAGAACAACGUGCCCGUGCAGC UAACGUCAAUCAGCGCACCUGAGCACAAGUUCGAGGGCCUGACCCA GAUCUUCCAGAAGGCCUACGAACACGAACAGCACAUCUCCGAGAGC AUCAACAAUAUUGUGGAUCACGCUAUCAAGUCCAAGGACCACGCUA CCUUCAACUUCCUGCAGUGGUACGUGGCCGAGCAACAUGAGGAGGA GGUGCUGUUCAAGGACAUCCUGGACAAGAUCGAGCUGAUCGGUAAU GAGAAUCACGGCCUGUACCUGGCCGACCAGUACGUGAAGGGCAUCG CCAAGAGCCGGAAGUCAGGCUCA	64

3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGG CCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGC	4
Corresponding amino acid sequence	MYSMQLASCVTLTLVLLVNSQPNITNLCPFGEVFNATRFASVYAWN RKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSF VIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVG GNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPL QSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKGGGSGGDIIK LLNEQVNKEMQSSNLYMSMSSWCYTHSLDGAGLFLFDHAAEEYEHA KKLIIFLNENNVPVQLTSISAPEHKFEGLTQIFQKAYEHEQHISES INNIVDHAIKSKDHATFNFLQWYVAEQHEEEVLFKDILDKIELIGN ENHGLYLADQYVKGIAKSRKSGS	65
PolyA tail	100 nt	

Lumazine Synthetase

In some embodiments, a lumazine synthetase domain is used as a scaffold domain. Lumazine synthetase is an enzyme responsible for the penultimate catalytic step in the biosynthesis of riboflavin in a variety of organisms, including archaea, bacteria, fungi, plants, and eubacteria. Lumazine synthetase is composed of homooligomers, which vary in size and subunit number, including pentamers, decamers, and icosahedral sixty-mers, depending on its species of origin. The lumazine synthetase monomer is 150 amino acids long and includes beta-sheets with flanking, tandem alpha-helices. Different quaternary structures have been reported for lumazine synthetase, illustrating its morphological versatility: from homopentamers up to symmetrical assemblies of twelve (12) pentamers forming capsids of 150 Å diameter. Presentation of antigens on the surface of lumazine synthetase results in a high local concentration of antigens displayed in an ordered array. Such repetitive structures enable the cross-linking of B-cell receptors and result in strong immune responses through an avidity effect.

An mRNA provided herein, in some embodiments, encodes an RBD linked to a lumazine synthetase domain, for example, through a glycine-serine (e.g., GGS). Other linkers may be used.

In other embodiments, an mRNA provided herein encodes an S1 domain of an S protein linked to a lumazine synthetase domain, for example, through a glycine-serine (e.g., GGS) linker. As indicated elsewhere herein, other linkers may be used.

Non-limiting examples of SARS-CoV-2 antigens linked to a foldon domain and the mRNA encoding them are provided in **Tables 14A** and **14B** below.

Table 14A. Antigens Linked to a Lumazine Synthetase Domain

Name	SEQ ID NO:	
SARS-CoV-2 Soluble S1 Linked to Lumazine Synthetase C-terminus (LS-S1)	10	11
SARS-CoV-2 Soluble S1 Linked to Lumazine Synthetase N-Terminus (S1-LS)	13	14

Name	SEQ ID NO:	
SARS-CoV-2 RBD Linked to Lumazine Synthetase C Terminus (LS-RBD)	67	68
SARS-CoV-2 RBD Linked to Lumazine Synthetase N Terminus (RBD-LS)	70	71

Table 14B. Antigens Linked to a Lumazine Synthetase Domain

SARS-CoV-2 Soluble S1 Linked to Lumazine Synthetase C-terminus (LS-S1)		
SEQ ID NO: 9 consists of from 5' end to 3' end: 5' UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 10 and 3' UTR SEQ ID NO: 4.		9
Chemistry	1-methylpseudouridine	
Cap	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCG GCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGGGCAUCCUGCCCAGCCCUGGCAUGCCCGCUCUGCUGAGCCUGG UGAGCCUGCUGAGCGUGCUGCUGAUGGGCUGCGUGGCUGAGACCGG CAUGCAGAUCUACGAGGGCAAGCUGACCGCAGAGGGCCUGCGGUUC GGCAUCGUGGCCAGCCGCGCCAACCACGCUCUGGUGGACCGGCUUG UGGAGGGCGCUAUCGACGCCAUCGUGAGACACGGCGGCCGGGAAGA GGACAUCACCCUGGUGCGGGUGUGCGGCAGCUGGGAGAUUCCCGUC GCCGCCGGAGAACUGGCCCGGAAGGAGGACAUCGACGCCGUGAUCG CCAUCGGCGUGCUGUGCAGAGGCGCCACGCCCAGCUUCGACUACAU CGCCAGCGAGGUGAGCAAGGGCCUGGCCGACCUGAGCCUGGAGCUG CGGAAGCCCAUCACCUUCGGCGUGAUCACCGCCGACACCCUGGAGC AGGCCAUCGAGGCCGCAGGCACCUGCCACGGCAACAAGGGCUGGGA AGCCGCCCUGUGCGCCAUCGAGAUGGCCAACCUGUUCAAGAGCCUG CGGGGCGGAAGUGGAGGCUCUGGUGGCAGCGGAGGAUCUGGCGGCG GCACCACCCGGACCCAGCUGCCACCAGCCUACACCAACAGCUUCAC CCGGGGCGUCUACUACCCCGACAAGGUGUUCCGGAGCAGCGUCCUG CACAGCACCCAGGACCUGUUCCUGCCCUUCUUCAGCAACGUGACCU GGUUCCACGCCAUCCACGUGAGCGGCACCAACGGCACCAAGCGGUU CGACAACCCCCUGCUGCCCUUCAACGACGGCGUGUACUUCGCCAGC ACCGAGAAGAGCAACAUCAUCCGGGGCUGGAUCUUCGGCACCACCC UGGACAGCAAGACCCAGAGCCUGCUGAUCGUGAAUAACGCCACCAA CGUGGUGAUCAAGGUGUGCGAGUUCCAGUUCUGCAACGACCCCUUC CUGGGCGUGUACUACCACAAGAACAACAAGAGCUGGAUGGAGAGCG AGUUCCGGGUGUACAGCAGCGCCAACAACUGCACCUUCGAGUACGU GAGCCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAGGGCAACUUC AAGAACCUGCGGGAGUUCGUGUUCAAGAACAUCGACGGCUACUUCA AGAUCUACAGCAAGCACACCCCAAUCAACCUGGUGCGGGAUCUGCC CCAGGGCUUCUCAGCCCUGGAGCCCCUGGUGGACCUGCCCAUCGGC AUCAACAUCACCCGGUUCCAGACCCUGCUGGCCCUGCACCGGAGCU ACCUGACCCCAGGCGACAGCAGCAGCGGGUGGACAGCAGGCGCGGC UGCUUACUACGUGGGCUACCUGCAGCCCCGGACCUUCCUGCUGAAG UACAACGAGAACGGCACCAUCACCGACGCCGUGGACUGCGCCCUGG ACCCUCUGAGCGAGACCAAGUGCACCCUGAAGAGCUUCACCGUGGA GAAGGGCAUCUACCAGACCAGCAACUUCCGGGUGCAGCCCACCGAG AGCAUCGUGCGGUUCCCCAACAUCACCAACCUGUGCCCCUUCGGCG AGGUGUUCAACGCCACCCGGUUCGCCAGCGUGUACGCCUGGAACCG GAAGCGGAUCAGCAACUGCGUGGCCGACUACAGCGUGCUGUACAAC AGCGCCAGCUUCAGCACCUUCAAGUGCUACGGCGUGAGCCCCACCA AGCUGAACGACCUGUGCUUCACCAACGUGUACGCCGACAGCUUCGU GAUCCGUGGCGACGAGGUGCGGCAGAUCGCACCCGGCCAGACAGGC AAGAUCGCCGACUACAACUACAAGCUGCCCGACGACUUCACCGGCU GCGUGAUCGCCUGGAACAGCAACAACCUCGACAGCAAGGUGGGCGG CAACUACAACUACCUGUACCGGCUGUUCCGGAAGAGCAACCUGAAG CCCUUCGAGCGGGACAUCAGCACCGAGAUCUACCAAGCCGGCUCCA	10

	CCCCUUGCAACGGCGUGGAGGGCUUCAACUGCUACUUCCCUCUGCA GAGCUACGGCUUCCAGCCCACCAACGGCGUGGGCUACCAGCCCUAC CGGGUGGUGGUGCUGAGCUUCGAGCUGCUGCACGCCCCAGCCACCG UGUGUGGCCCCAAGAAGAGCACCAACCUGGUGAAGAACAAGUGCGU GAACUUCAACUUCAACGGCCUUACCGGCACCGGCGUGCUGACCGAG AGCAACAAGAAAUUCCUGCCCUUUCAGCAGUUCGGCCGGGACAUCG CCGACACCACCGACGCUGUGCGGGAUCCCCAGACCCUGGAGAUCCU GGACAUCACCCCUUGCAGCUUCGGCGGCGUGAGCGUGAUCACCCCA GGCACCAACACCAGCAACCAGGUGGCCGUGCUGUACCAGGACGUGA ACUGCACCGAGGUGCCCGUGGCCAUCCACGCCGACCAGCUGACACC CACCUGGCGGGUCUACAGCACCGGCAGCAACGUGUUCCAGACCCGG GCCGGUUGCCUGAUCGGCGCCGAGCACGUGAACAACAGCUACGAGU GCGACAUCCCCAUCGGCGCCGGCAUCUGUGCCAGCUACCAGACCCA GACCAAUUCA	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGG CCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGC	4
Corresponding amino acid sequence	MGILPSPGMPALLSLVSLLSVLLMGCVAETGMQIYEGKLTAEGLRF GIVASRANHALVDRLVEGAIDAIVRHGGREEDITLVRVCGSWEIPV AAGELARKEDIDAVIAIGVLCRGATPSFDYIASEVSKGLADLSLEL RKPITFGVITADTLEQAIEAAGTCHGNKGWEAALCAIEMANLFKSL RGGSGGSGGSGGSGGGTTRTQLPPAYTNSFTRGVYYPDKVFRSSVL HSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFAS TEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPF LGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNF KNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIG INITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLK YNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQPTE SIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYN SASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTG KIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLK PFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPY RVVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTE SNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITP GTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTR AGCLIGAEHVNNSYECDIPIGAGICASYQTQTNS	11
PolyA tail	100 nt	
SARS-CoV-2 Soluble S1 Linked to Lumazine Synthetase N-Terminus (S1-LS)		
SEQ ID NO: 12 consists of from 5' end to 3' end: 5' UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 13 and 3' UTR SEQ ID NO: 4.		12
Chemistry	1-methylpseudouridine	
Cap	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCG GCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGUGAGCAGCCAGUGCG UGAACCUGACCACCCGGACCCAGCUGCCACCAGCCUACACCAACAG CUUCACCCGGGGCGUCUACUACCCCGACAAGGUGUUCCGGAGCAGC GUCCUGCACAGCACCCAGGACCUGUUCCUGCCCUUCUUCAGCAACG UGACCUGGUUCCACGCCAUCCACGUGAGCGGCACCAACGGCACCAA GCGGUUCGACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUC GCCAGCACCGAGAAGAGCAACAUCAUCCGGGGCUGGAUCUUCGGCA CCACCCUGGACAGCAAGACCCAGAGCCUGCUGAUCGUGAAUAACGC CACCAACGUGGUGAUCAAGGUGUGCGAGUUCCAGUUCUGCAACGAC CCCUUCCUGGGCGUGUACUACCACAAGAACAACAAGAGCUGGAUGG AGAGCGAGUUCCGGGUGUACAGCAGCGCCAACAACUGCACCUUCGA GUACGUGAGCCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAGGGC AACUUCAAGAACCUGCGGGAGUUCGUGUUCAAGAACAUCGACGGCU ACUUCAAGAUCUACAGCAAGCACACCCCAAUCAACCUGGUGCGGGA UCUGCCCCAGGGCUUCUCAGCCCUGGAGCCCCUGGUGGACCUGCCC AUCGGCAUCAACAUCACCCGGUUCCAGACCCUGCUGGCCCUGCACC	13

	GGAGCUACCUGACCCCAGGCGACAGCAGCAGCGGGUGGACAGCAGG CGCGGCUGCUUACUACGUGGGCUACCUGCAGCCCCGGACCUUCCUG CUGAAGUACAACGAGAACGGCACCAUCACCGACGCCGUGGACUGCG CCCUGGACCCUCUGAGCGAGACCAAGUGCACCCUGAAGAGCUUCAC CGUGGAGAAGGGCAUCUACCAGACCAGCAACUUCCGGGUGCAGCCC ACCGAGAGCAUCGUGCGGUUCCCCAACAUCACCAACCUGUGCCCCU UCGGCGAGGUGUUCAACGCCACCCGGUUCGCCAGCGUGUACGCCUG GAACCGGAAGCGGAUCAGCAACUGCGUGGCCGACUACAGCGUGCUG UACAACAGCGCCAGCUUCAGCACCUUCAAGUGCUACGGCGUGAGCC CCACCAAGCUGAACGACCUGUGCUUCACCAACGUGUACGCCGACAG CUUCGUGAUCCGUGGCGACGAGGUGCGGCAGAUCGCACCCGGCCAG ACAGGCAAGAUCGCCGACUACAACUACAAGCUGCCCGACGACUUCA CCGGCUGCGUGAUCGCCUGGAACAGCAACAACCUCGACAGCAAGGU GGGCGGCAACUACAACUACCUGUACCGGCUGUUCCGGAAGAGCAAC CUGAAGCCCUUCGAGCGGGACAUCAGCACCGAGAUCUACCAAGCCG GCUCCACCCCUUGCAACGGCGUGGAGGGCUUCAACUGCUACUUCCC UCUGCAGAGCUACGGCUUCCAGCCCACCAACGGCGUGGGCUACCAG CCCUACCGGGUGGUGGUGCUGAGCUUCGAGCUGCUGCACGCCCCAG CCACCGUGUGUGGCCCCAAGAAGAGCACCAACCUGGUGAAGAACAA GUGCGUGAACUUCAACUUCAACGGCCUUACCGGCACCGGCGUGCUG ACCGAGAGCAACAAGAAAUUCCUGCCCUUUCAGCAGUUCGGCCGGG ACAUCGCCGACACCACCGACGCUGUGCGGGAUCCCCAGACCCUGGA GAUCCUGGACAUCACCCCUUGCAGCUUCGGCGGCGUGAGCGUGAUC ACCCCAGGCACCAACACCAGCAACCAGGUGGCCGUGCUGUACCAGG ACGUGAACUGCACCGAGGUGCCCGUGGCCAUCCACGCCGACCAGCU GACACCCACCUGGCGGGUCUACAGCACCGGCAGCAACGUGUUCCAG ACCCGGGCCGGUUGCCUGAUCGGCGCCGAGCACGUGAACAACAGCU ACGAGUGCGACAUCCCCAUCGGCGCCGGCAUCUGUGCCAGCUACCA GACCCAGACCAAUUCAGGAGGAGGCUCCGGAGGCGGUAGCGCUGAG ACCGGCAUGCAGAUCUACGAGGGCAAGCUGACCGCAGAGGGCCUGC GGUUCGGCAUCGUGGCCAGCCGCGCCAACCACGCUCUGGUGGACCG GCUUGUGGAGGGCGCUAUCGACGCCAUCGUGAGACACGGCGGCCGG GAAGAGGACAUCACCCUGGUGCGGGUGUGCGGCAGCUGGGAGAUUC CCGUCGCCGCCGGAGAACUGGCCCGGAAGGAGGACAUCGACGCCGU GAUCGCCAUCGGCGUGCUGUGCAGAGGCGCCACGCCCAGCUUCGAC UACAUCGCCAGCGAGGUGAGCAAGGGCCUGGCCGACCUGAGCCUGG AGCUGCGGAAGCCCAUCACCUUCGGCGUGAUCACCGCCGACACCCU GGAGCAGGCCAUCGAGGCCGCAGGCACCUGCCACGGCAACAAGGGC UGGGAAGCCGCCCUGUGCGCCAUCGAGAUGGCCAACCUGUUCAAGA GCCUGCGG	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGG CCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGC	4
Corresponding amino acid sequence	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSS VLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYF ASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCND PFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQG NFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLP IGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFL LKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQP TESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVL YNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQ TGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSN LKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQ PYRVVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVL TESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVI TPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQ TRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSgggsgggsAE TGMQIYEGKLTAEGLRFGIVASRANHALVDRLVEGAIDAIVRHGGR EEDITLVRVCGSWEIPVAAGELARKEDIDAVIAIGVLCRGATPSFD YIASEVSKGLADLSLELRKPITFGVITADTLEQAIEAAGTCHGNKG WEAALCAIEMANLFKSLR	14

PolyA tail	100 nt	
SARS-CoV-2 RBD Linked to Lumazine Synthetase C Terminus (LS-RBD)		
SEQ ID NO: 66 consists of from 5' end to 3' end: 5' UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 67 and 3' UTR SEQ ID NO: 4.		66
Chemistry	1-methylpseudouridine	
Cap	7mG(5')ppp(5')NImpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCG GCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGGGCAUCCUGCCCAGCCCUGGCAUGCCCGCUCUGCUGAGCCUGG UGAGCCUGCUGAGCGUGCUGCUGAUGGGCUGCGUGGCUGAGACCGG CAUGCAGAUCUACGAGGGCAAGCUGACCGCAGAGGGCCUGCGGUUC GGCAUCGUGGCCAGCCGCGCCAACCACGCUCUGGUGGACCGGCUUG UGGAGGGCGCUAUCGACGCCAUCGUGAGACACGGCGGCCGGGAAGA GGACAUCACCCUGGUGCGGGUGUGCGGCAGCUGGGAGAUUCCCGUC GCCGCCGGAGAACUGGCCCGGAAGGAGGACAUCGACGCCGUGAUCG CCAUCGGCGUGCUGUGCAGAGGCGCCACGCCCAGCUUCGACUACAU CGCCAGCGAGGUGAGCAAGGGCCUGGCCGACCUGAGCCUGGAGCUG CGGAAGCCCAUCACCUUCGGCGUGAUCACCGCCGACACCCUGGAGC AGGCCAUCGAGGCCGCAGGCACCUGCCACGGCAACAAGGGCUGGGA AGCCGCCCUGUGCGCCAUCGAGAUGGCCAACCUGUUCAAGAGCCUG CGGGGCGGAAGUGGAGGCUCUGGUGGCAGCGGAGGAUCUGGCGGCG GCCAGCCCAACAUCACCAACCUGUGCCCCUUCGGCGAGGUGUUCAA CGCCACCCGGUUCGCCAGCGUGUACGCCUGGAACCGGAAGCGGAUC AGCAACUGCGUGGCCGACUACAGCGUGCUGUACAACAGCGCCAGCU UCAGCACCUUCAAGUGCUACGGCGUGAGCCCCACCAAGCUGAACGA CCUGUGCUUCACCAACGUGUACGCCGACAGCUUCGUGAUCCGUGGC GACGAGGUGCGGCAGAUCGCACCCGGCCAGACAGGCAAGAUCGCCG ACUACAACUACAAGCUGCCCGACGACUUCACCGGCUGCGUGAUCGC CUGGAACAGCAACAACCUCGACAGCAAGGUGGGCGGCAACUACAAC UACCUGUACCGGCUGUUCCGGAAGAGCAACCUGAAGCCCUUCGAGC GGGACAUCAGCACCGAGAUCUACCAAGCCGGCUCCACCCCUUGCAA CGGCGUGGAGGGCUUCAACUGCUACUUCCCCUCUGCAGAGCUACGGC UUCCAGCCCACCAACGGCGUGGGCUACCAGCCCUACCGGGUGGUGG UGCUGAGCUUCGAGCUGCUGCACGCCCCAGCCACCGUGUGUGGCCC CAAG	67
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGG CCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGC	4
Corresponding amino acid sequence	MGILPSPGMPALLSLVSLLSVLLMGCVAETGMQIYEGKLTAEGLRF GIVASRANHALVDRLVEGAIDAIVRHGGREEDITLVRVCGSWEIPV AAGELARKEDIDAVIAIGVLCRGATPSFDYIASEVSKGLADLSLEL RKPITFGVITADTLEQAIEAAGTCHGNKGWEAALCAIEMANLFKSL RGGSGGSGGSGGSGGGQPNITNLCPFGEVFNATRFASVYAWNRKRI SNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRG DEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYN YLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYG FQPTNGVGYQPYRVVVLSFELLHAPATVCGPK	68
PolyA tail	100 nt	
SARS-CoV-2 RBD Linked to Lumazine Synthetase N Terminus (RBD-LS)		
SEQ ID NO: 69 consists of from 5' end to 3' end: 5' UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 70 and 3' UTR SEQ ID NO: 4.		69
Chemistry	1-methylpseudouridine	
Cap	7mG(5')ppp(5')NImpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCG GCGCCGCCACC	2
ORF of mRNA Construct	AUGUACAGCAUGCAGCUGGCUAGCUGCGUGACCCUGACCCUGGUGC UGCUGGUGAACAGCCAGCCCAACAUCACCAACCUGUGCCCCUUCGG	70

(excluding the stop codon)	CGAGGUGUUCAACGCCACCCGGUUCGCCAGCGUGUACGCCUGGAAC CGGAAGCGGAUCAGCAACUGCGUGGCCGACUACAGCGUGCUGUACA ACAGCGCCAGCUUCAGCACCUUCAAGUGCUACGGCGUGAGCCCCAC CAAGCUGAACGACCUGUGCUUCACCAACGUGUACGCCGACAGCUUC GUGAUCCGUGGCGACGAGGUGCGGCAGAUCGCACCCGGCCAGACAG GCAAGAUCGCCGACUACAACUACAAGCUGCCCGACGACUUCACCGG CUGCGUGAUCGCCUGGAACAGCAACAACCUCGACAGCAAGGUGGGC GGCAACUACAACUACCUGUACCGGCUGUUCCGGAAGAGCAACCUGA AGCCCUUCGAGCGGGACAUCAGCACCGAGAUCUACCAAGCCGGCUC CACCCCUUGCAACGGCGUGGAGGGCUUCAACUGCUACUUCCCUCUG CAGAGCUACGGCUUCCAGCCCACCAACGGCGUGGGCUACCAGCCCU ACCGGGUGGUGGUGCUGAGCUUCGAGCUGCUGCACGCCCCAGCCAC CGUGUGUGGCCCCAAGGGAGGAGGCUCCGGAGGCGGUAGCGCUGAG ACCGGCAUGCAGAUCUACGAGGGCAAGCUGACCGCAGAGGGCCUGC GGUUCGGCAUCGUGGCCAGCCGCGCCAACCACGCUCUGGUGGACCG GCUUGUGGAGGGCGCUAUCGACGCCAUCGUGAGACACGGCGGCCGG GAAGAGGACAUCACCCUGGUGCGGGUGUGCGGCAGCUGGGAGAUUC CCGUCGCCGCCGGAGAACUGGCCCGGAAGGAGGACAUCGACGCCGU GAUCGCCAUCGGCGUGCUGUGCAGAGGCGCCACGCCCAGCUUCGAC UACAUCGCCAGCGAGGUGAGCAAGGGCCUGGCCGACCUGAGCCUGG AGCUGCGGAAGCCCAUCACCUUCGGCGUGAUCACCGCCGACACCCU GGAGCAGGCCAUCGAGGCCGCAGGCACCUGCCACGGCAACAAGGGC UGGGAAGCCGCCCUGUGCGCCAUCGAGAUGGCCAACCUGUUCAAGA GCCUGCGG	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGG CCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGC	4
Corresponding amino acid sequence	MYSMQLASCVTLTLVLLVNSQPNITNLCPFGEVFNATRFASVYAWN RKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSF VIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVG GNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPL QSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKgggsgggsAE TGMQIYEGKLTAEGLRFGIVASRANHALVDRLVEGAIDAIVRHGGR EEDITLVRVCGSWEIPVAAGELARKEDIDAVIAIGVLCRGATPSFD YIASEVSKGLADLSLELRKPITFGVITADTLEQAIEAAGTCHGNKG WEAALCAIEMANLFKSLR	71
PolyA tail	100 nt	

Foldon

In some embodiments, a foldon domain is used as a scaffold domain. The C-terminal domain of T4 fibritin (foldon) is obligatory for the formation of the fibritin trimer structure and can be used as an artificial trimerization domain (see, e.g., Meier S. et al. Journal of Molecular Biology 2004 Dec 3; 344(4): 1051-1069; Tao Y et al. Structure 1997 Jun 15; 5(6):789-98). When fused to the S protein ectodomain, a foldon domain promotes correct trimerization of the S protein, thus avoiding misfolding of the protein. Such a process resulting in production of the prefusion conformation of the S protein results in increased expression, conformational homogeneity, and elicitation of potent neutralizing antibody responses.

Without being bound by theory, it is thought that this configuration would result in the foldon being largely immunogenically silent on the intracellular region of the protein. Non-limiting examples of SARS-CoV-2 antigens linked to a foldon domain and the mRNA encoding them are provided in **Tables 15A** and **15B** below.

Table 15A. Antigens Linked to a Foldon Domain

Name	SEQ ID NO:	
	mRNA ORF	Protein
SARS-CoV-2 NTD Linked to Foldon Domain	43	44
SARS-CoV-2 NTD Linked to Transmembrane Domain and Foldon Domain	49	50
SARS-CoV-2 RBD Linked to Foldon Domain	73	74
SARS-CoV-2 RBD Linked to Foldon Domain and Transmembrane Domain (RBD-FD-TM)	79	80
SARS-CoV-2 RBD Linked to Transmembrane Domain and Foldon Domain (RBD-TM-FD)	82	83
SARS-CoV-2 NTD-RBD Linked to Foldon Domain and Transmembrane Domain (NTD-RBD-FD-TM)	100	101
SARS-CoV-2 NTD-RBD Linked to Transmembrane Domain and Foldon Domain (NTD-RBD-TM-FD)	103	104
SARS-CoV-2 NTD-RBD Linked to Foldon Domain	112	113

Table 15B. Antigens Linked to a Foldon Domain

SARS-CoV-2 NTD Linked to Foldon Domain		
SEQ ID NO: 42 consists of from 5' end to 3' end: 5' UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 43 and 3' UTR SEQ ID NO: 4.		42
Chemistry	1-methylpseudouridine	
Cap	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCG GCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGUGAGCAGCCAGUGCG UGAACCUGACCACCCGGACCCAGCUGCCACCAGCCUACACCAACAG CUUCACCCGGGGCGUCUACUACCCCGACAAGGUGUUCCGGAGCAGC GUCCUGCACAGCACCCAGGACCUGUUCCUGCCCUUCUUCAGCAACG UGACCUGGUUCCACGCCAUCCACGUGAGCGGCACCAACGGCACCAA GCGGUUCGACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUC GCCAGCACCGAGAAGAGCAACAUCAUCCGGGGCUGGAUCUUCGGCA CCACCCUGGACAGCAAGACCCAGAGCCUGCUGAUCGUGAAUAACGC CACCAACGUGGUGAUCAAGGUGUGCGAGUUCCAGUUCUGCAACGAC CCCUUCCUGGGCGUGUACUACCACAAGAACAACAAGAGCUGGAUGG AGAGCGAGUUCCGGGUGUACAGCAGCGCCAACAACUGCACCUUCGA GUACGUGAGCCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAGGGC AACUUCAAGAACCUGCGGGAGUUCGUGUUCAAGAACAUCGACGGCU ACUUCAAGAUCUACAGCAAGCACACCCCAAUCAACCUGGUGCGGGA UCUGCCCCAGGGCUUCUCAGCCCUGGAGCCCCUGGUGGACCUGCCC AUCGGCAUCAACAUCACCCGGUUCCAGACCCUGCUGGCCCUGCACC GGAGCUACCUGACCCCAGGCGACAGCAGCAGCGGGUGGACAGCAGG CGCGGCUGCUUACUACGUGGGCUACCUGCAGCCCCGGACCUUCCUG CUGAAGUACAACGAGAACGGCACCAUCACCGACGCCGUGGACUCUG GCGGAGGCGGCAGCGCCAUCGGCGGCUACAUCCCCGAGGCCCCUAG AGACGGCCAGGCCUACGUGCGGAAGGACGGCGAGUGGGUGCUGCUG AGCACCUUCCUGGGC	43
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGG CCUCCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGC	4
Corresponding amino acid sequence	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSS VLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYF	44

	ASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCND PFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQG NFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLP IGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFL LKYNENGTITDAVDsggggSAIGGYIPEAPRDGQAYVRKDGEWVLL STFLG	
PolyA tail	100 nt	
SARS-CoV-2 NTD Linked to Transmembrane Domain and Foldon Domain		
SEQ ID NO: 48 consists of from 5' end to 3' end: 5' UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 49 and 3' UTR SEQ ID NO: 4.		48
Chemistry	1-methylpseudouridine	
Cap	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCG GCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGUGAGCAGCCAGUGCG UGAACCUGACCACCCGGACCCAGCUGCCACCAGCCUACACCAACAG CUUCACCCGGGGCGUCUACUACCCCGACAAGGUGUUCCGGAGCAGC GUCCUGCACAGCACCCAGGACCUGUUCCUGCCCUUCUUCAGCAACG UGACCUGGUUCCACGCCAUCCACGUGAGCGGCACCAACGGCACCAA GCGGUUCGACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUC GCCAGCACCGAGAAGAGCAACAUCAUCCGGGGCUGGAUCUUCGGCA CCACCCUGGACAGCAAGACCCAGAGCCUGCUGAUCGUGAAUAACGC CACCAACGUGGUGAUCAAGGUGUGCGAGUUCCAGUUCUGCAACGAC CCCUUCCUGGGCGUGUACUACCACAAGAACAACAAGAGCUGGAUGG AGAGCGAGUUCCGGGUGUACAGCAGCGCCAACAACUGCACCUUCGA GUACGUGAGCCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAGGGC AACUUCAAGAACCUGCGGGAGUUCGUGUUCAAGAACAUCGACGGCU ACUUCAAGAUCUACAGCAAGCACACCCCAAUCAACCUGGUGCGGGA UCUGCCCCAGGGCUUCUCAGCCCUGGAGCCCCUGGUGGACCUGCCC AUCGGCAUCAACAUCACCCGGUUCCAGACCCUGCUGGCCCUGCACC GGAGCUACCUGACCCCAGGCGACAGCAGCAGCGGGUGGACAGCAGG CGCGGCUGCUUACUACGUGGGCUACCUGCAGCCCCGGACCUUCCUG CUGAAGUACAACGAGAACGGCACCAUCACCGACGCCGUGGACUCUG GCGGAGGCAGCAUCCUGGCCAUCUACAGCACCGUGGCCAGCAGCCU GGUGCUGCUGGUGAGCCUGGGCGCCAUCAGCUUCGGCGGAGGCAGC GCCAUCGGCGGCUACAUCCCCGAGGCCCCUAGAGACGGCCAGGCCU ACGUGCGGAAGGACGGCGAGUGGGUGCUGCUGAGCACCUUCCUGGG CAAG	49
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGG CCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGC	4
Corresponding amino acid sequence	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSS VLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYF ASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCND PFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQG NFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLP IGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFL LKYNENGTITDAVDsgggsilaiystvasslvllvslgaisfgggS AIGGYIPEAPRDGQAYVRKDGEWVLLSTFLGk	50
PolyA tail	100 nt	
SARS-CoV-2 RBD Linked to Foldon Domain		
SEQ ID NO: 72 consists of from 5' end to 3' end: 5' UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 73 and 3' UTR SEQ ID NO: 4.		72
Chemistry	1-methylpseudouridine	
Cap	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCG GCGCCGCCACC	2

ORF of mRNA Construct (excluding the stop codon)	AUGUACAGCAUGCAGCUGGCUAGCUGCGUGACCCUGACCCUGGUGC UGCUGGUGAACAGCCAGCCCAACAUCACCAACCUGUGCCCCUUCGG CGAGGUGUUCAACGCCACCCGGUUCGCCAGCGUGUACGCCUGGAAC CGGAAGCGGAUCAGCAACUGCGUGGCCGACUACAGCGUGCUGUACA ACAGCGCCAGCUUCAGCACCUUCAAGUGCUACGGCGUGAGCCCCAC CAAGCUGAACGACCUGUGCUUCACCAACGUGUACGCCGACAGCUUC GUGAUCCGUGGCGACGAGGUGCGGCAGAUCGCACCCGGCCAGACAG GCAAGAUCGCCGACUACAACUACAAGCUGCCCGACGACUUCACCGG CUGCGUGAUCGCCUGGAACAGCAACAACCUCGACAGCAAGGUGGGC GGCAACUACAACUACCUGUACCGGCUGUUCCGGAAGAGCAACCUGA AGCCCUUCGAGCGGGACAUCAGCACCGAGAUCUACCAAGCCGGCUC CACCCCUUGCAACGGCGUGGAGGGCUUCAACUGCUACUUCCCUCUG CAGAGCUACGGCUUCCAGCCCACCAACGGCGUGGGCUACCAGCCCU ACCGGGUGGUGGUGCUGAGCUUCGAGCUGCUGCACGCCCCAGCCAC CGUGUGUGGCCCCAAGUCUGGCGGAGGCGGCAGCGCCAUCGGCGGC UACAUCCCCGAGGCCCCUAGAGACGGCCAGGCCUACGUGCGGAAGG ACGGCGAGUGGGUGCUGCUGAGCACCUUCCUGGGC	73
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGG CCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGC	4
Corresponding amino acid sequence	mysmqlascvtltlvllvnsQPNITNLCPFGEVFNATRFASVYAWN RKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSF VIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVG GNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPL QSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKsggggSAIGG YIPEAPRDGQAYVRKDGEWVLLSTFLG	74
PolyA tail	100 nt	
SARS-CoV-2 RBD Linked to Foldon Domain and Transmembrane Domain (RBD-FD-TM)		
SEQ ID NO: 78 consists of from 5' end to 3' end: 5' UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 79 and 3' UTR SEQ ID NO: 4.		78
Chemistry	1-methylpseudouridine	
Cap	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCG GCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUACAGCAUGCAGCUGGCUAGCUGCGUGACCCUGACCCUGGUGC UGCUGGUGAACAGCCAGCCCAACAUCACCAACCUGUGCCCCUUCGG CGAGGUGUUCAACGCCACCCGGUUCGCCAGCGUGUACGCCUGGAAC CGGAAGCGGAUCAGCAACUGCGUGGCCGACUACAGCGUGCUGUACA ACAGCGCCAGCUUCAGCACCUUCAAGUGCUACGGCGUGAGCCCCAC CAAGCUGAACGACCUGUGCUUCACCAACGUGUACGCCGACAGCUUC GUGAUCCGUGGCGACGAGGUGCGGCAGAUCGCACCCGGCCAGACAG GCAAGAUCGCCGACUACAACUACAAGCUGCCCGACGACUUCACCGG CUGCGUGAUCGCCCUGGAACAGCAACAACCUCGACAGCAAGGUGGGC GGCAACUACAACUACCUGUACCGGCUGUUCCGGAAGAGCAACCUGA AGCCCUUCGAGCGGGACAUCAGCACCGAGAUCUACCAAGCCGGCUC CACCCCUUGCAACGGCGUGGAGGGCUUCAACUGCUACUUCCCUCUG CAGAGCUACGGCUUCCAGCCCACCAACGGCGUGGGCUACCAGCCCU ACCGGGUGGUGGUGCUGAGCUUCGAGCUGCUGCACGCCCCAGCCAC CGUGUGUGGCCCCAAGUCUGGCGGAGGCGGCAGCGCCAUCGGCGGC UACAUCCCCGAGGCCCCUAGAGACGGCCAGGCCUACGUGCGGAAGG ACGGCGAGUGGGUGCUGCUGAGCACCUUCCUGGGCGGAGGCAGCAU CCUGGCCAUCUACAGCACCGUGGCCAGCAGCCUGGUGCUGCUGGUG AGCCUGGGCGCCAUCAGCUUC	79
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGG CCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGC	4
Corresponding amino acid sequence	mysmqlascvtltlvllvnsQPNITNLCPFGEVFNATRFASVYAWN RKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSF VIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVG GNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPL	80

	QSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKsggggSAIGG YIPEAPRDGQAYVRKDGEWVLLSTFLGggsilaiystvasslvllv slgaisf	
PolyA tail	100 nt	
SARS-CoV-2 RBD Linked to Transmembrane Domain and Foldon Domain (RBD-TM-FD)		
SEQ ID NO: 81 consists of from 5' end to 3' end: 5' UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 82 and 3' UTR SEQ ID NO: 4.		81
Chemistry	1-methylpseudouridine	
Cap	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCG GCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUACAGCAUGCAGCUGGCUAGCUGCGUGACCCUGACCCUGGUGC UGCUGGUGAACAGCCAGCCCAACAUCACCAACCUGUGCCCCUUCGG CGAGGUGUUCAACGCCACCCGGUUCGCCAGCGUGUACGCCUGGAAC CGGAAGCGGAUCAGCAACUGCGUGGCCGACUACAGCGUGCUGUACA ACAGCGCCAGCUUCAGCACCUUCAAGUGCUACGGCGUGAGCCCCAC CAAGCUGAACGACCUGUGCUUCACCAACGUGUACGCCGACAGCUUC GUGAUCCGUGGCGACGAGGUGCGGCAGAUCGCACCCGGCCAGACAG GCAAGAUCGCCGACUACAACUACAAGCUGCCCGACGACUUCACCGG CUGCGUGAUCGCCUGGAACAGCAACAACCUCGACAGCAAGGUGGGC GGCAACUACAACUACCUGUACCGGCUGUUCCGGAAGAGCAACCUGA AGCCCUUCGAGCGGGACAUCAGCACCGAGAUCUACCAAGCCGGCUC CACCCCUUGCAACGGCGUGGAGGGCUUCAACUGCUACUUCCCUCUG CAGAGCUACGGCUUCCAGCCCACCAACGGCGUGGGCUACCAGCCCU ACCGGGUGGUGGUGCUGAGCUUCGAGCUGCUGCACGCCCCAGCCAC CGUGUGUGGCCCCAAGUCUGGCGGAGGCAGCAUCCUGGCCAUCUAC AGCACCGUGGCCAGCAGCCUGGUGCUGCUGGUGAGCCUGGGCGCCA UCAGCUUCGGCGGAGGCAGCGCCAUCGGCGGCUACAUCCCCGAGGC CCCUAGAGACGGCCAGGCCUACGUGCGGAAGGACGGCGAGUGGGUG CUGCUGAGCACCUUCCUGGGCAAG	82
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGG CCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGC	4
Corresponding amino acid sequence	mysmqlascvtltlvllvnsQPNITNLCPFGEVFNATRFASVYAWN RKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSF VIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVG GNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPL QSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKsgggsilaiy stvasslvllvslgaisfgggSAIGGYIPEAPRDGQAYVRKDGEWV LLSTFLGk	83
PolyA tail	100 nt	
SARS-CoV-2 NTD-RBD Linked to Foldon Domain and Transmembrane Domain (NTD-RBD-FD-TM)		
SEQ ID NO: 99 consists of from 5' end to 3' end: 5' UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 100 and 3' UTR SEQ ID NO: 4.		99
Chemistry	1-methylpseudouridine	
Cap	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCG GCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGUGAGCAGCCAGUGCG UGAACCUGACCACCCGGACCCAGCUGCCACCAGCCUACACCAACAG CUUCACCCGGGGCGUCUACUACCCCGACAAGGUGUUCCGGAGCAGC GUCCUGCACAGCACCCAGGACCUGUUCCUGCCCUUCUUCAGCAACG UGACCUGGUUCCACGCCAUCCACGUGAGCGGCACCAACGGCACCAA GCGGUUCGACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUC GCCAGCACCGAGAAGAGCAACAUCAUCCGGGGCUGGAUCUUCGGCA CCACCCUGGACAGCAAGACCCAGAGCCUGCUGAUCGUGAAUAACGC CACCAACGUGGUGAUCAAGGUGUGCGAGUUCCAGUUCUGCAACGAC	100

	CCCUUCCUGGGCGUGUACUACCACAAGAACAACAAGAGCUGGAUGG AGAGCGAGUUCCGGGUGUACAGCAGCGCCAACAACUGCACCUUCGA GUACGUGAGCCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAGGGC AACUUCAAGAACCUGCGGGAGUUCGUGUUCAAGAACAUCGACGGCU ACUUCAAGAUCUACAGCAAGCACACCCCAAUCAACCUGGUGCGGGA UCUGCCCCAGGGCUUCUCAGCCCUGGAGCCCCUGGUGGACCUGCCC AUCGGCAUCAACAUCACCCGGUUCCAGACCCUGCUGGCCCUGCACC GGAGCUACCUGACCCCAGGCGACAGCAGCAGCGGGUGGACAGCAGG CGCGGCUGCUUACUACGUGGGCUACCUGCAGCCCCGGACCUUCCUG CUGAAGUACAACGAGAACGGCACCAUCACCGACGCCGUGGACGGAG GCGGAUCGGGAGGCGGACCCAACAUCACCAACCUGUGCCCCUUCGG CGAGGUGUUCAACGCCACCCGGUUCGCCAGCGUGUACGCCUGGAAC CGGAAGCGGAUCAGCAACUGCGUGGCCGACUACAGCGUGCUGUACA ACAGCGCCAGCUUCAGCACCUUCAAGUGCUACGGCGUGAGCCCCAC CAAGCUGAACGACCUGUGCUUCACCAACGUGUACGCCGACAGCUUC GUGAUCCGUGGCGACGAGGUGCGGCAGAUCGCACCCGGCCAGACAG GCAAGAUCGCCGACUACAACUACAAGCUGCCCGACGACUUCACCGG CUGCGUGAUCGCCUGGAACAGCAACAACCUCGACAGCAAGGUGGGC GGCAACUACAACUACCUGUACCGGCUGUUCCGGAAGAGCAACCUGA AGCCCUUCGAGCGGGACAUCAGCACCGAGAUCUACCAAGCCGGCUC CACCCCUUGCAACGGCGUGGAGGGCUUCAACUGCUACUUCCCUCUG CAGAGCUACGGCUUCCAGCCCACCAACGGCGUGGGCUACCAGCCCU ACCGGGUGGUGGUGCUGAGCUUCGAGCUGCUGCACGCCCCAGCCAC CGUGUGUGGCCCCAAGUCUGGCGGAGGCGGCAGCGCCAUCGGCGGC UACAUCCCCGAGGCCCCUAGAGACGGCCAGGCCUACGUGCGGAAGG ACGGCGAGUGGGUGCUGCUGAGCACCUUCCUGGGCGGAGGCAGCAU CCUGGCCAUCUACAGCACCGUGGCCAGCAGCCUGGUGCUGCUGGUG AGCCUGGGCGCCAUCAGCUUC	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGG CCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGC	4
Corresponding amino acid sequence	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSS VLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYF ASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCND PFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQG NFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLP IGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFL LKYNENGTITDAVDgggsgggPNITNLCPFGEVFNATRFASVYAWN RKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSF VIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVG GNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPL QSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKsggggSAIGG YIPEAPRDGQAYVRKDGEWVLLSTFLGggsilaiystvasslvllv slgaisf	101
PolyA tail	100 nt	
SARS-CoV-2 NTD-RBD Linked to Transmembrane Domain and Foldon Domain (NTD-RBD-TM-FD)		
SEQ ID NO: 102 consists of from 5' end to 3' end: 5' UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 103 and 3' UTR SEQ ID NO: 4.		102
Chemistry	1-methylpseudouridine	
Cap	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCG GCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGUGAGCAGCCAGUGCG UGAACCUGACCACCCGGACCCAGCUGCCACCAGCCUACACCAACAG CUUCACCCGGGGCGUCUACUACCCCGACAAGGUGUUCCGGAGCAGC GUCCUGCACAGCACCCAGGACCUGUUCCUGCCCUUCUUCAGCAACG UGACCUGGUUCCACGCCAUCCACGUGAGCGGCACCAACGGCACCAA GCGGUUCGACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUC GCCAGCACCGAGAAGAGCAACAUCAUCCGGGGCUGGAUCUUCGGCA CCACCCUGGACAGCAAGACCCAGAGCCUGCUGAUCGUGAAUAACGC	103

	CACCAACGUGGUGAUCAAGGUGUGCGAGUUCCAGUUCUGCAACGAC CCCUUCCUGGGCGUGUACUACCACAAGAACAACAAGAGCUGGAUGG AGAGCGAGUUCCGGGUGUACAGCAGCGCCAACAACUGCACCUUCGA GUACGUGAGCCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAGGGC AACUUCAAGAACCUGCGGGAGUUCGUGUUCAAGAACAUCGACGGCU ACUUCAAGAUCUACAGCAAGCACACCCCAAUCAACCUGGUGCGGGA UCUGCCCCAGGGCUUCUCAGCCCUGGAGCCCCUGGUGGACCUGCCC AUCGGCAUCAACAUCACCCGGUUCCAGACCCUGCUGGCCCUGCACC GGAGCUACCUGACCCCAGGCGACAGCAGCAGCGGGUGGACAGCAGG CGCGGCUGCUUACUACGUGGGCUACCUGCAGCCCCGGACCUUCCUG CUGAAGUACAACGAGAACGGCACCAUCACCGACGCCGUGGACGGAG GCGGAUCGGGAGGCGGACCCAACAUCACCAACCUGUGCCCCUUCGG CGAGGUGUUCAACGCCACCCGGUUCGCCAGCGUGUACGCCUGGAAC CGGAAGCGGAUCAGCAACUGCGUGGCCGACUACAGCGUGCUGUACA ACAGCGCCAGCUUCAGCACCUUCAAGUGCUACGGCGUGAGCCCCAC CAAGCUGAACGACCUGUGCUUCACCAACGUGUACGCCGACAGCUUC GUGAUCCGUGGCGACGAGGUGCGGCAGAUCGGCACCCGGCCAGACAG GCAAGAUCGCCGACUACAACUACAAGCUGCCCGACGACUUCACCGG CUGCGUGAUCGCCUGGAACAGCAACAACCUCGACAGCAAGGUGGGC GGCAACUACAACUACCUGUACCGGCUGUUCCGGAAGAGCAACCUGA AGCCCUUCGAGCGGGACAUCAGCACCGAGAUCUACCAAGCCGGCUC CACCCCUUGCAACGGCGUGGAGGGCUUCAACUGCUACUUCCCUCUG CAGAGCUACGGCUUCCAGCCCACCAACGGCGUGGGCUACCAGCCCU ACCGGGUGGUGGUGCUGAGCUUCGAGCUGCUGCACGCCCCAGCCAC CGUGUGUGGCCCCAAGUCUGGCGGAGGCAGCAUCCUGGCCAUCUAC AGCACCGUGGCCAGCAGCCUGGUGCUGCUGGUGAGCCUGGGCGCCA UCAGCUUCGGCGGAGGCAGCGCCAUCGGCGGCUACAUCCCCGAGGC CCCUAGAGACGGCCAGGCCUACGUGCGGAAGGACGGCGAGUGGGUG CUGCUGAGCACCUUCCUGGGCAAG	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGG CCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGC	4
Corresponding amino acid sequence	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSS VLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYF ASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCND PFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQG NFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLP IGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFL LKYNENGTITDAVDgggsgggPNITNLCPFGEVFNATRFASVYAWN RKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSF VIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVG GNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPL QSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKsgggsilaiy stvasslvllvslgaisfgggSAIGGYIPEAPRDGQAYVRKDGEWV LLSTFLGk	104
PolyA tail	100 nt	
SARS-CoV-2 NTD-RBD Linked to Foldon Domain		
SEQ ID NO: 111 consists of from 5' end to 3' end: 5' UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 112 and 3' UTR SEQ ID NO: 4.		111
Chemistry	1-methylpseudouridine	
Cap	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCG GCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGUGAGCAGCCAGUGCG UGAACCUGACCACCCGGACCCAGCUGCCACCAGCCUACACCAACAG CUUCACCCGGGGCGUCUACUACCCCGACAAGGUGUUCCGGAGCAGC GUCCUGCACAGCACCCAGGACCUGUUCCUGCCCUUCUUCAGCAACG UGACCUGGUUCCACGCCAUCCACGUGAGCGGCACCAACGGCACCAA GCGGUUCGACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUC GCCAGCACCGAGAAGAGCAACAUCAUCCGGGGCUGGAUCUUCGGCA	112

	CCACCCUGGACAGCAAGACCCAGAGCCUGCUGAUCGUGAAUAACGC CACCAACGUGGUGAUCAAGGUGUGCGAGUUCCAGUUCUGCAACGAC CCCUUCCUGGGCGUGUACUACCACAAGAACAACAAGAGCUGGAUGG AGAGCGAGUUCCGGGUGUACAGCAGCGCCAACAACUGCACCUUCGA GUACGUGAGCCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAGGGC AACUUCAAGAACCUGCGGGAGUUCGUGUUCAAGAACAUCGACGGCU ACUUCAAGAUCUACAGCAAGCACACCCCAAUCAACCUGGUGCGGGA UCUGCCCCAGGGCUUCUCAGCCCUGGAGCCCCUGGUGGACCUGCCC AUCGGCAUCAACAUCACCCGGUUCCAGACCCUGCUGGCCCUGCACC GGAGCUACCUGACCCCAGGCGACAGCAGCAGCGGGUGGACAGCAGG CGCGGCUGCUUACUACGUGGGCUACCUGCAGCCCCGGACCUUCCUG CUGAAGUACAACGAGAACGGCACCAUCACCGACGCCGUGGACGGAG GCGGAUCGGGAGGCGGACCCAACAUCACCAACCUGUGCCCCUUCGG CGAGGUGUUCAACGCCACCCGGUUCGCCAGCGUGUACGCCUGGAAC CGGAAGCGGAUCAGCAACUGCGUGGCCGACUACAGCGUGCUGUACA ACAGCGCCAGCUUCAGCACCUUCAAGUGCUACGGCGUGAGCCCCAC CAAGCUGAACGACCUGUGCUUCACCAACGUGUACGCCGACAGCUUC GUGAUCCGUGGCGACGAGGUGCGGCAGAUCGCACCCGGCCAGACAG GCAAGAUCGCCGACUACAACUACAAGCUGCCCGACGACUUCACCGG CUGCGUGAUCGCCUGGAACAGCAACAACCUCGACAGCAAGGUGGGC GGCAACUACAACUACCUGUACCGGCUGUUCCGGAAGAGCAACCUGA AGCCCUUCGAGCGGGACAUCAGCACCGAGAUCUACCAAGCCGGCUC CACCCCUUGCAACGGCGUGGAGGGCUUCAACUGCUACUUCCCUCUG CAGAGCUACGGCUUCCAGCCCACCAACGGCGUGGGCUACCAGCCCU ACCGGGUGGUGGUGCUGAGCUUCGAGCUGCUGCACGCCCCAGCCAC CGUGUGUGGCCCCAAGUCUGGCGGAGGCGGCAGCGCCAUCGGCGGC UACAUCCCCGAGGCCCCUAGAGACGGCCAGGCCUACGUGCGGAAGG ACGGCGAGUGGGUGCUGCUGAGCACCUUCCUGGGC	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGG CCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGC	4
Corresponding amino acid sequence	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSS VLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYF ASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCND PFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQG NFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLP IGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFL LKYNENGTITDAVDgggsgggPNITNLCPFGEVFNATRFASVYAWN RKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSF VIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVG GNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPL QSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKsggggSAIGG YIPEAPRDGQAYVRKDGEWVLLSTFLG	113
PolyA tail	100 nt	

Encapsulin

In some embodiments, an encapsulin domain is used as a scaffold domain. Encapsulin is a protein cage nanoparticle isolated from the thermophile Thermotoga maritima. Encapsulin is assembled from 60 copies of identical 31 kDa monomers having a thin and icosahedral T = 1 symmetric cage structure with interior and exterior diameters of 20 and 24 nm, respectively (Sutter M. et al. Nat Struct Mol Biol. 2008; 15: 939-947). Although the exact function of encapsulin in T. maritima is not clearly understood yet, its crystal structure has been recently solved and its function was postulated as a cellular compartment that encapsulates proteins such as DyP (Dye decolorizing peroxidase) and Flp (Ferritin like protein), which are involved in

oxidative stress responses 30 (Rahmanpour R. et al. FEBS J. 2013; 280: 2097-2104). The use of encapsulin for nanoparticle construction enables both the display of protein antigen on the surface of the nanoparticle, and the enclosure of cargo such as mRNA within the nanoparticle itself. Previous encapsulin nanoparticle-based vaccines have elicited strong immune responses to both surface displayed antigen and cargo protein itself (Lagoutte P. et al. Vaccine. 2018; 36(25): 3622–3628).

An mRNA provided herein, in some embodiments, encodes an S protein domain (e.g., S1, S2, RBD, and/or NTD) linked to an encapsulin domain.

Fusion Proteins

In some embodiments, a composition of the present disclosure includes an mRNA encoding an antigenic fusion protein. Thus, the encoded antigen or antigens may include two or more proteins (e.g., protein and/or protein fragment) joined together. Alternatively, the protein to which a protein antigen is fused does not promote a strong immune response to itself, but rather to the coronavirus antigen. Antigenic fusion proteins, in some embodiments, retain the functional property from each original protein.

In some embodiments, a fusion protein comprises a receptor binding domain from a SARS-CoV-2 Spike protein.

In some embodiments, a fusion protein comprises an N-terminal domain from a SARS-CoV-2 Spike protein

In some embodiments, a fusion protein comprises a transmembrane domain. The transmembrane domain may, in some embodiments, be from a virus that is not SARS-CoV-2. For example, the transmembrane domain may be from an influenza hemagglutinin transmembrane domain, which has been demonstrated to effectively anchor proteins at the cell surface.

Variants

In some embodiments, the compositions of the present disclosure include RNA that encodes a coronavirus antigen variant. Antigen variants or other polypeptide variants refers to molecules that differ in their amino acid sequence from a wild-type, native, or reference sequence. The antigen/polypeptide variants may possess substitutions, deletions, and/or insertions at certain positions within the amino acid sequence, as compared to a native or reference sequence. Ordinarily, variants possess at least 50% identity to a wild-type, native or

reference sequence. In some embodiments, variants share at least 80%, or at least 90% identity with a wild-type, native, or reference sequence.

Variant antigens/polypeptides encoded by nucleic acids of the disclosure may contain amino acid changes that confer any of a number of desirable properties, e.g., that enhance their immunogenicity, enhance their expression, and/or improve their stability or PK/PD properties in a subject. Variant antigens/polypeptides can be made using routine mutagenesis techniques and assayed as appropriate to determine whether they possess the desired property. Assays to determine expression levels and immunogenicity are well known in the art and exemplary such assays are set forth in the Examples section. Similarly, PK/PD properties of a protein variant can be measured using art recognized techniques, e.g., by determining expression of antigens in a vaccinated subject over time and/or by looking at the durability of the induced immune response. The stability of protein(s) encoded by a variant nucleic acid may be measured by assaying thermal stability or stability upon urea denaturation or may be measured using in silico prediction. Methods for such experiments and in silico determinations are known in the art.

In some embodiments, a composition comprises an mRNA or an mRNA ORF that comprises a nucleotide sequence of any one of the sequences provided herein (see, e.g., Sequence Listing), or comprises a nucleotide sequence at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical to a nucleotide sequence of any one of the sequences provided herein.

The term "identity" refers to a relationship between the sequences of two or more polypeptides (e.g. antigens) or polynucleotides (nucleic acids), as determined by comparing the sequences. Identity also refers to the degree of sequence relatedness between or among sequences as determined by the number of matches between strings of two or more amino acid residues or nucleic acid residues. Identity measures the percent of identical matches between the smaller of two or more sequences with gap alignments (if any) addressed by a particular mathematical model or computer program (e.g., "algorithms"). Identity of related antigens or nucleic acids can be readily calculated by known methods. "Percent (%) identity" as it applies to polypeptide or polynucleotide sequences is defined as the percentage of residues (amino acid residues or nucleic acid residues) in the candidate amino acid or nucleic acid sequence that are identical with the residues in the amino acid sequence or nucleic acid sequence of a second sequence after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent identity. Methods and computer programs for the alignment are well known in the art. It is understood that identity depends on a calculation of percent identity but may differ in value due to gaps and penalties introduced in the calculation. Generally, variants of a particular

polynucleotide or polypeptide (e.g., antigen) have at least 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% but less than 100% sequence identity to that particular reference polynucleotide or polypeptide as determined by sequence alignment programs and parameters described herein and known to those skilled in the art. Such tools for alignment include those of the BLAST suite (Stephen F. Altschul, et al (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nucleic Acids Res. 25:3389-3402). Another popular local alignment technique is based on the Smith-Waterman algorithm (Smith, T.F. & Waterman, M.S. (1981) "Identification of common molecular subsequences." J. Mol. Biol. 147:195-197). A general global alignment technique based on dynamic programming is the Needleman–Wunsch algorithm (Needleman, S.B. & Wunsch, C.D. (1970) "A general method applicable to the search for similarities in the amino acid sequences of two proteins." J. Mol. Biol. 48:443-453). More recently a Fast Optimal Global Sequence Alignment Algorithm (FOGSAA) has been developed that purportedly produces global alignment of nucleotide and protein sequences faster than other optimal global alignment methods, including the Needleman–Wunsch algorithm.

As such, polynucleotides encoding peptides or polypeptides containing substitutions, insertions and/or additions, deletions and covalent modifications with respect to reference sequences, particularly the polypeptide (e.g., antigen) sequences disclosed herein, are included within the scope of this disclosure. For example, sequence tags or amino acids, such as one or more lysines, can be added to peptide sequences (e.g., at the N-terminal or C-terminal ends). Sequence tags can be used for peptide detection, purification or localization. Lysines can be used to increase peptide solubility or to allow for biotinylation. Alternatively, amino acid residues located at the carboxy and amino terminal regions of the amino acid sequence of a peptide or protein may optionally be deleted providing for truncated sequences. Certain amino acids (e.g., C-terminal or N-terminal residues) may alternatively be deleted depending on the use of the sequence, as for example, expression of the sequence as part of a larger sequence which is soluble or linked to a solid support. In some embodiments, sequences for (or encoding) signal sequences, termination sequences, transmembrane domains, linkers, multimerization domains (such as, e.g., foldon regions) and the like may be substituted with alternative sequences that achieve the same or a similar function. In some embodiments, cavities in the core of proteins can be filled to improve stability, e.g., by introducing larger amino acids. In other embodiments, buried hydrogen bond networks may be replaced with hydrophobic resides to improve stability. In yet other embodiments, glycosylation sites may be removed and replaced with appropriate residues. Such sequences are readily identifiable to one of skill in the art. It should also be

understood that some of the sequences provided herein contain sequence tags or terminal peptide sequences (e.g., at the N-terminal or C-terminal ends) that may be deleted, for example, prior to use in the preparation of an mRNA vaccine.

As recognized by those skilled in the art, protein fragments, functional protein domains, and homologous proteins are also considered to be within the scope of coronavirus antigens of interest. For example, provided herein is any protein fragment (meaning a polypeptide sequence at least one amino acid residue shorter than a reference antigen sequence but otherwise identical) of a reference protein, provided that the fragment is immunogenic and confers a protective immune response to the coronavirus. In addition to variants that are identical to the reference protein but are truncated, in some embodiments, an antigen includes 2, 3, 4, 5, 6, 7, 8, 9, 10, or more mutations, as shown in any of the sequences provided or referenced herein. Antigens/antigenic polypeptides can range in length from about 4, 6, or 8 amino acids to full length proteins.

Stabilizing Elements

Naturally-occurring eukaryotic mRNA molecules can contain stabilizing elements, including, but not limited to untranslated regions (UTR) at their 5′-end (5′ UTR) and/or at their 3′-end (3′ UTR), in addition to other structural features, such as a 5′-cap structure or a 3′-poly(A) tail. Both the 5′ UTR and the 3′ UTR are typically transcribed from the genomic DNA and are elements of the premature mRNA. Characteristic structural features of mature mRNA, such as the 5′-cap and the 3′-poly(A) tail are usually added to the transcribed (premature) mRNA during mRNA processing.

In some embodiments, a composition includes an mRNA having an open reading frame encoding at least one antigenic polypeptide having at least one modification, at least one 5′ terminal cap, and is formulated within a lipid nanoparticle. 5′-capping of polynucleotides may be completed concomitantly during the in vitro-transcription reaction using the following chemical RNA cap analogs to generate the 5′-guanosine cap structure according to manufacturer protocols: 3´-O-Me-m7G(5')ppp(5') G [the ARCA cap];G(5')ppp(5')A; G(5')ppp(5')G; m7G(5')ppp(5')A; m7G(5')ppp(5')G (New England BioLabs, Ipswich, MA). 5′-capping of modified RNA may be completed post-transcriptionally using a Vaccinia Virus Capping Enzyme to generate the “Cap 0” structure: m7G(5')ppp(5')G (New England BioLabs, Ipswich, MA). Cap 1 structure may be generated using both Vaccinia Virus Capping Enzyme and a 2′-O methyl-transferase to generate: m7G(5')ppp(5')G-2′-O-methyl. Cap 2 structure may be generated from the Cap 1 structure followed by the 2′-O-methylation of the 5′-antepenultimate nucleotide using a 2′-O methyl-

transferase. Cap 3 structure may be generated from the Cap 2 structure followed by the 2′-O-methylation of the 5′-preantepenultimate nucleotide using a 2′-O methyl-transferase. Enzymes may be derived from a recombinant source.

The 3′-poly(A) tail is typically a stretch of adenine nucleotides added to the 3′-end of the transcribed mRNA. It can, in some instances, comprise up to about 400 adenine nucleotides. In some embodiments, the length of the 3′-poly(A) tail may be an essential element with respect to the stability of the individual mRNA.

In some embodiments, a composition includes a stabilizing element. Stabilizing elements may include for instance a histone stem-loop. A stem-loop binding protein (SLBP), a 32 kDa protein has been identified. It is associated with the histone stem-loop at the 3'-end of the histone messages in both the nucleus and the cytoplasm. Its expression level is regulated by the cell cycle; it peaks during the S-phase, when histone mRNA levels are also elevated. The protein has been shown to be essential for efficient 3'-end processing of histone pre-mRNA by the U7 snRNP. SLBP continues to be associated with the stem-loop after processing, and then stimulates the translation of mature histone mRNAs into histone proteins in the cytoplasm. The RNA binding domain of SLBP is conserved through metazoa and protozoa; its binding to the histone stem-loop depends on the structure of the loop. The minimum binding site includes at least three nucleotides 5’ and two nucleotides 3′ relative to the stem-loop.

In some embodiments, an mRNA includes a coding region, at least one histone stem-loop, and optionally, a poly(A) sequence or polyadenylation signal. The poly(A) sequence or polyadenylation signal generally should enhance the expression level of the encoded protein. The encoded protein, in some embodiments, is not a histone protein, a reporter protein (e.g. Luciferase, GFP, EGFP, β-Galactosidase, EGFP), or a marker or selection protein (e.g. alpha-Globin, Galactokinase and Xanthine:guanine phosphoribosyl transferase (GPT)).

In some embodiments, an mRNA includes the combination of a poly(A) sequence or polyadenylation signal and at least one histone stem-loop, even though both represent alternative mechanisms in nature, acts synergistically to increase the protein expression beyond the level observed with either of the individual elements. The synergistic effect of the combination of poly(A) and at least one histone stem-loop does not depend on the order of the elements or the length of the poly(A) sequence.

In some embodiments, an mRNA does not include a histone downstream element (HDE). “Histone downstream element” (HDE) includes a purine-rich polynucleotide stretch of approximately 15 to 20 nucleotides 3′ of naturally occurring stem-loops, representing the binding

site for the U7 snRNA, which is involved in processing of histone pre-mRNA into mature histone mRNA. In some embodiments, the nucleic acid does not include an intron.

An mRNA may or may not contain an enhancer and/or promoter sequence, which may be modified or unmodified or which may be activated or inactivated. In some embodiments, the histone stem-loop is generally derived from histone genes and includes an intramolecular base pairing of two neighbored partially or entirely reverse complementary sequences separated by a spacer, consisting of a short sequence, which forms the loop of the structure. The unpaired loop region is typically unable to base pair with either of the stem loop elements. It occurs more often in RNA, as is a key component of many RNA secondary structures but may be present in single-stranded DNA as well. Stability of the stem-loop structure generally depends on the length, number of mismatches or bulges, and base composition of the paired region. In some embodiments, wobble base pairing (non-Watson-Crick base pairing) may result. In some embodiments, the at least one histone stem-loop sequence comprises a length of 15 to 45 nucleotides.

In some embodiments, an mRNA has one or more AU-rich sequences removed. These sequences, sometimes referred to as AURES are destabilizing sequences found in the 3'UTR. The AURES may be removed from the RNA vaccines. Alternatively, the AURES may remain in the RNA vaccine.

Signal Peptides

In some embodiments, a composition comprises an mRNA having an ORF that encodes a signal peptide fused to the coronavirus antigen. Signal peptides, comprising the N-terminal 15-60 amino acids of proteins, are typically needed for the translocation across the membrane on the secretory pathway and, thus, universally control the entry of most proteins both in eukaryotes and prokaryotes to the secretory pathway. In eukaryotes, the signal peptide of a nascent precursor protein (pre-protein) directs the ribosome to the rough endoplasmic reticulum (ER) membrane and initiates the transport of the growing peptide chain across it for processing. ER processing produces mature proteins, wherein the signal peptide is cleaved from precursor proteins, typically by a ER-resident signal peptidase of the host cell, or they remain uncleaved and function as a membrane anchor. A signal peptide may also facilitate the targeting of the protein to the cell membrane.

A signal peptide may have a length of 15-60 amino acids. For example, a signal peptide may have a length of 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, or

60 amino acids. In some embodiments, a signal peptide has a length of 20-60, 25-60, 30-60, 35-60, 40-60, 45- 60, 50-60, 55-60, 15-55, 20-55, 25-55, 30-55, 35-55, 40-55, 45-55, 50-55, 15-50, 20-50, 25-50, 30-50, 35-50, 40-50, 45-50, 15-45, 20-45, 25-45, 30-45, 35-45, 40-45, 15-40, 20-40, 25-40, 30-40, 35-40, 15-35, 20-35, 25-35, 30-35, 15-30, 20-30, 25-30, 15-25, 20-25, or 15-20 amino acids.

Signal peptides from heterologous genes (which regulate expression of genes other than coronavirus antigens in nature) are known in the art and can be tested for desired properties and then incorporated into a nucleic acid of the disclosure.

Sequence Optimization

In some embodiments, an ORF encoding an antigen of the disclosure is codon optimized. Codon optimization methods are known in the art. For example, an ORF of any one or more of the sequences provided herein may be codon optimized. Codon optimization, in some embodiments, may be used to match codon frequencies in target and host organisms to ensure proper folding; bias GC content to increase mRNA stability or reduce secondary structures; minimize tandem repeat codons or base runs that may impair gene construction or expression; customize transcriptional and translational control regions; insert or remove protein trafficking sequences; remove/add post translation modification sites in encoded protein (e.g., glycosylation sites); add, remove or shuffle protein domains; insert or delete restriction sites; modify ribosome binding sites and mRNA degradation sites; adjust translational rates to allow the various domains of the protein to fold properly; or reduce or eliminate problem secondary structures within the polynucleotide. Codon optimization tools, algorithms and services are known in the art – non-limiting examples include services from GeneArt (Life Technologies), DNA2.0 (Menlo Park CA) and/or proprietary methods. In some embodiments, the open reading frame (ORF) sequence is optimized using optimization algorithms.

In some embodiments, a codon optimized sequence shares less than 95% sequence identity to a naturally-occurring or wild-type sequence ORF (e.g., a naturally-occurring or wild-type mRNA sequence encoding a coronavirus antigen). In some embodiments, a codon optimized sequence shares less than 90% sequence identity to a naturally-occurring or wild-type sequence (e.g., a naturally-occurring or wild-type mRNA sequence encoding a coronavirus antigen). In some embodiments, a codon optimized sequence shares less than 85% sequence identity to a naturally-occurring or wild-type sequence (e.g., a naturally-occurring or wild-type mRNA sequence encoding a coronavirus antigen). In some embodiments, a codon optimized sequence shares less than 80% sequence identity to a naturally-occurring or wild-type sequence

(e.g., a naturally-occurring or wild-type mRNA sequence encoding a coronavirus antigen). In some embodiments, a codon optimized sequence shares less than 75% sequence identity to a naturally-occurring or wild-type sequence (e.g., a naturally-occurring or wild-type mRNA sequence encoding a coronavirus antigen).

In some embodiments, a codon optimized sequence shares between 65% and 85% (e.g., between about 67% and about 85% or between about 67% and about 80%) sequence identity to a naturally-occurring or wild-type sequence (e.g., a naturally-occurring or wild-type mRNA sequence encoding a coronavirus antigen). In some embodiments, a codon optimized sequence shares between 65% and 75% or about 80% sequence identity to a naturally-occurring or wild-type sequence (e.g., a naturally-occurring or wild-type mRNA sequence encoding a coronavirus antigen).

In some embodiments, a codon-optimized sequence encodes an antigen that is as immunogenic as, or more immunogenic than (e.g., at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 100%, or at least 200% more), than a coronavirus antigen encoded by a non-codon-optimized sequence.

When transfected into mammalian host cells, the modified mRNAs have a stability of between 12-18 hours, or greater than 18 hours, e.g., 24, 36, 48, 60, 72, or greater than 72 hours and are capable of being expressed by the mammalian host cells.

In some embodiments, a codon optimized RNA may be one in which the levels of G/C are enhanced. The G/C-content of nucleic acid molecules (e.g., mRNA) may influence the stability of the RNA. RNA having an increased amount of guanine (G) and/or cytosine (C) residues may be functionally more stable than mRNA containing a large amount of adenine (A) and thymine (T) or uracil (U) nucleotides. As an example, WO02/098443 discloses a pharmaceutical composition containing an mRNA stabilized by sequence modifications in the translated region. Due to the degeneracy of the genetic code, the modifications work by substituting existing codons for those that promote greater RNA stability without changing the resulting amino acid. The approach is limited to coding regions of the RNA.

Chemically Unmodified Nucleotides

In some embodiments, an mRNA is not chemically modified and comprises the standard ribonucleotides consisting of adenosine, guanosine, cytosine and uridine. In some embodiments, nucleotides and nucleosides of the present disclosure comprise standard nucleoside residues such as those present in transcribed RNA (e.g. A, G, C, or U). In some embodiments, nucleotides and

nucleosides of the present disclosure comprise standard deoxyribonucleosides such as those present in DNA (e.g. dA, dG, dC, or dT).

Chemical Modifications

The compositions of the present disclosure comprise, in some embodiments, an mRNA having an open reading frame encoding a coronavirus antigen, wherein the nucleic acid comprises nucleotides and/or nucleosides that can be standard (unmodified) or modified as is known in the art. In some embodiments, nucleotides and nucleosides of the present disclosure comprise modified nucleotides or nucleosides. Such modified nucleotides and nucleosides can be naturally-occurring modified nucleotides and nucleosides or non-naturally occurring modified nucleotides and nucleosides. Such modifications can include those at the sugar, backbone, or nucleobase portion of the nucleotide and/or nucleoside as are recognized in the art.

In some embodiments, a naturally-occurring modified nucleotide or nucleotide of the disclosure is one as is generally known or recognized in the art. Non-limiting examples of such naturally occurring modified nucleotides and nucleotides can be found, inter alia, in the widely recognized MODOMICS database.

In some embodiments, a non-naturally occurring modified nucleotide or nucleoside of the disclosure is one as is generally known or recognized in the art. Non-limiting examples of such non-naturally occurring modified nucleotides and nucleosides can be found, inter alia, in published US application Nos. PCT/US2012/058519; PCT/US2013/075177; PCT/US2014/058897; PCT/US2014/058891; PCT/US2014/070413; PCT/US2015/36773; PCT/US2015/36759; PCT/US2015/36771; or PCT/IB2017/051367 all of which are incorporated by reference herein.

Hence, nucleic acids of the disclosure (e.g., DNA nucleic acids and RNA nucleic acids, such as mRNA nucleic acids) can comprise standard nucleotides and nucleosides, naturally-occurring nucleotides and nucleosides, non-naturally-occurring nucleotides and nucleosides, or any combination thereof.

Nucleic acids of the disclosure (e.g., DNA nucleic acids and RNA nucleic acids, such as mRNA nucleic acids), in some embodiments, comprise various (more than one) different types of standard and/or modified nucleotides and nucleosides. In some embodiments, a particular region of a nucleic acid contains one, two or more (optionally different) types of standard and/or modified nucleotides and nucleosides.

In some embodiments, a modified RNA nucleic acid (e.g., a modified mRNA nucleic acid), introduced to a cell or organism, exhibits reduced degradation in the cell or organism,

respectively, relative to an unmodified nucleic acid comprising standard nucleotides and nucleosides.

In some embodiments, a modified RNA nucleic acid (e.g., a modified mRNA nucleic acid), introduced into a cell or organism, may exhibit reduced immunogenicity in the cell or organism, respectively (e.g., a reduced innate response) relative to an unmodified nucleic acid comprising standard nucleotides and nucleosides.

Nucleic acids (e.g., RNA nucleic acids, such as mRNA nucleic acids), in some embodiments, comprise non-natural modified nucleotides that are introduced during synthesis or post-synthesis of the nucleic acids to achieve desired functions or properties. The modifications may be present on internucleotide linkages, purine or pyrimidine bases, or sugars. The modification may be introduced with chemical synthesis or with a polymerase enzyme at the terminal of a chain or anywhere else in the chain. Any of the regions of a nucleic acid may be chemically modified.

The present disclosure provides for modified nucleosides and nucleotides of a nucleic acid (e.g., RNA nucleic acids, such as mRNA nucleic acids). A "nucleoside" refers to a compound containing a sugar molecule (e.g., a pentose or ribose) or a derivative thereof in combination with an organic base (e.g., a purine or pyrimidine) or a derivative thereof (also referred to herein as "nucleobase"). A "nucleotide" refers to a nucleoside, including a phosphate group. Modified nucleotides may by synthesized by any useful method, such as, for example, chemically, enzymatically, or recombinantly, to include one or more modified or non-natural nucleosides. Nucleic acids can comprise a region or regions of linked nucleosides. Such regions may have variable backbone linkages. The linkages can be standard phosphodiester linkages, in which case the nucleic acids would comprise regions of nucleotides.

Modified nucleotide base pairing encompasses not only the standard adenosine-thymine, adenosine-uracil, or guanosine-cytosine base pairs, but also base pairs formed between nucleotides and/or modified nucleotides comprising non-standard or modified bases, wherein the arrangement of hydrogen bond donors and hydrogen bond acceptors permits hydrogen bonding between a non-standard base and a standard base or between two complementary non-standard base structures, such as, for example, in those nucleic acids having at least one chemical modification. One example of such non-standard base pairing is the base pairing between the modified nucleotide inosine and adenine, cytosine or uracil. Any combination of base/sugar or linker may be incorporated into nucleic acids of the present disclosure.

In some embodiments, modified nucleobases in nucleic acids (e.g., RNA nucleic acids, such as mRNA nucleic acids) comprise 1-methyl-pseudouridine (m1ψ), 1-ethyl-pseudouridine

(e1ψ), 5-methoxy-uridine (mo5U), 5-methyl-cytidine (m5C), and/or pseudouridine (ψ). In some embodiments, modified nucleobases in nucleic acids (e.g., RNA nucleic acids, such as mRNA nucleic acids) comprise 5-methoxymethyl uridine, 5-methylthio uridine, 1-methoxymethyl pseudouridine, 5-methyl cytidine, and/or 5-methoxy cytidine. In some embodiments, the polyribonucleotide includes a combination of at least two (e.g., 2, 3, 4 or more) of any of the aforementioned modified nucleobases, including but not limited to chemical modifications.

In some embodiments, an mRNA of the disclosure comprises 1-methyl-pseudouridine (m1ψ) substitutions at one or more or all uridine positions of the nucleic acid.

In some embodiments, an mRNA of the disclosure comprises 1-methyl-pseudouridine (m1ψ) substitutions at one or more or all uridine positions of the nucleic acid and 5-methyl cytidine substitutions at one or more or all cytidine positions of the nucleic acid.

In some embodiments, an mRNA of the disclosure comprises pseudouridine (ψ) substitutions at one or more or all uridine positions of the nucleic acid.

In some embodiments, an mRNA of the disclosure comprises pseudouridine (ψ) substitutions at one or more or all uridine positions of the nucleic acid and 5-methyl cytidine substitutions at one or more or all cytidine positions of the nucleic acid.

In some embodiments, an mRNA of the disclosure comprises uridine at one or more or all uridine positions of the nucleic acid.

In some embodiments, mRNAs are uniformly modified (e.g., fully modified, modified throughout the entire sequence) for a particular modification. For example, a nucleic acid can be uniformly modified with 1-methyl-pseudouridine, meaning that all uridine residues in the mRNA sequence are replaced with 1-methyl-pseudouridine. Similarly, a nucleic acid can be uniformly modified for any type of nucleoside residue present in the sequence by replacement with a modified residue such as those set forth above.

The nucleic acids of the present disclosure may be partially or fully modified along the entire length of the molecule. For example, one or more or all or a given type of nucleotide (e.g., purine or pyrimidine, or any one or more or all of A, G, U, C) may be uniformly modified in a nucleic acid of the disclosure, or in a predetermined sequence region thereof (e.g., in the mRNA including or excluding the poly(A) tail). In some embodiments, all nucleotides X in a nucleic acid of the present disclosure (or in a sequence region thereof) are modified nucleotides, wherein X may be any one of nucleotides A, G, U, C, or any one of the combinations A+G, A+U, A+C, G+U, G+C, U+C, A+G+U, A+G+C, G+U+C or A+G+C.

The nucleic acid may contain from about 1% to about 100% modified nucleotides (either in relation to overall nucleotide content, or in relation to one or more types of nucleotide, i.e., any

one or more of A, G, U or C) or any intervening percentage (e.g., from 1% to 20%, from 1% to 25%, from 1% to 50%, from 1% to 60%, from 1% to 70%, from 1% to 80%, from 1% to 90%, from 1% to 95%, from 10% to 20%, from 10% to 25%, from 10% to 50%, from 10% to 60%, from 10% to 70%, from 10% to 80%, from 10% to 90%, from 10% to 95%, from 10% to 100%, from 20% to 25%, from 20% to 50%, from 20% to 60%, from 20% to 70%, from 20% to 80%, from 20% to 90%, from 20% to 95%, from 20% to 100%, from 50% to 60%, from 50% to 70%, from 50% to 80%, from 50% to 90%, from 50% to 95%, from 50% to 100%, from 70% to 80%, from 70% to 90%, from 70% to 95%, from 70% to 100%, from 80% to 90%, from 80% to 95%, from 80% to 100%, from 90% to 95%, from 90% to 100%, and from 95% to 100%). It will be understood that any remaining percentage is accounted for by the presence of unmodified A, G, U, or C.

The mRNAs may contain at a minimum 1% and at maximum 100% modified nucleotides, or any intervening percentage, such as at least 5% modified nucleotides, at least 10% modified nucleotides, at least 25% modified nucleotides, at least 50% modified nucleotides, at least 80% modified nucleotides, or at least 90% modified nucleotides. For example, the nucleic acids may contain a modified pyrimidine such as a modified uracil or cytosine. In some embodiments, at least 5%, at least 10%, at least 25%, at least 50%, at least 80%, at least 90% or 100% of the uracil in the nucleic acid is replaced with a modified uracil (e.g., a 5-substituted uracil). The modified uracil can be replaced by a compound having a single unique structure or can be replaced by a plurality of compounds having different structures (e.g., 2, 3, 4 or more unique structures). In some embodiments, at least 5%, at least 10%, at least 25%, at least 50%, at least 80%, at least 90% or 100% of the cytosine in the nucleic acid is replaced with a modified cytosine (e.g., a 5-substituted cytosine). The modified cytosine can be replaced by a compound having a single unique structure or can be replaced by a plurality of compounds having different structures (e.g., 2, 3, 4 or more unique structures).

Untranslated Regions (UTRs)

The mRNAs of the present disclosure may comprise one or more regions or parts which act or function as an untranslated region. Where mRNAs are designed to encode at least one antigen of interest, the nucleic may comprise one or more of these untranslated regions (UTRs). Wild-type untranslated regions of a nucleic acid are transcribed but not translated. In mRNA, the 5′ UTR starts at the transcription start site and continues to the start codon but does not include the start codon; whereas, the 3′ UTR starts immediately following the stop codon and continues until the transcriptional termination signal. There is growing body of evidence about the

regulatory roles played by the UTRs in terms of stability of the nucleic acid molecule and translation. The regulatory features of a UTR can be incorporated into the polynucleotides of the present disclosure to, among other things, enhance the stability of the molecule. The specific features can also be incorporated to ensure controlled down-regulation of the transcript in case they are misdirected to undesired organs sites. A variety of 5'UTR and 3'UTR sequences are known and available in the art.

A 5' UTR is region of an mRNA that is directly upstream (5') from the start codon (the first codon of an mRNA transcript translated by a ribosome). A 5' UTR does not encode a protein (is non-coding). Natural 5'UTRs have features that play roles in translation initiation. They harbor signatures like Kozak sequences which are commonly known to be involved in the process by which the ribosome initiates translation of many genes. Kozak sequences have the consensus CCR(A/G)CCAUGG (SEQ ID NO: 128), where R is a purine (adenine or guanine) three bases upstream of the start codon (AUG), which is followed by another 'G'.5'UTR also have been known to form secondary structures which are involved in elongation factor binding.

In some embodiments of the disclosure, a 5' UTR is a heterologous UTR, i.e., is a UTR found in nature associated with a different ORF. In another embodiment, a 5' UTR is a synthetic UTR, i.e., does not occur in nature. Synthetic UTRs include UTRs that have been mutated to improve their properties, e.g., which increase gene expression as well as those which are completely synthetic. Exemplary 5' UTRs include Xenopus or human derived a-globin or b-globin (8278063; 9012219), human cytochrome b-245 a polypeptide, and hydroxysteroid (17b) dehydrogenase, and Tobacco etch virus (US8278063, 9012219). CMV immediate-early 1 (IE1) gene (US20140206753, WO2013/185069), the sequence GGGAUCCUACC (SEQ ID NO: 129) (WO2014144196) may also be used. In another embodiment, 5' UTR of a TOP gene is a 5' UTR of a TOP gene lacking the 5' TOP motif (the oligopyrimidine tract) (e.g., WO/2015101414, WO2015101415, WO/2015/062738, WO2015024667, WO2015024667; 5' UTR element derived from ribosomal protein Large 32 (L32) gene (WO/2015101414, WO2015101415, WO/2015/062738), 5' UTR element derived from the 5'UTR of an hydroxysteroid (17-β) dehydrogenase 4 gene (HSD17B4) (WO2015024667), or a 5' UTR element derived from the 5' UTR of ATP5A1 (WO2015024667) can be used. In some embodiments, an internal ribosome entry site (IRES) is used instead of a 5' UTR.

In some embodiments, a 5' UTR of the present disclosure comprises a sequence selected from SEQ ID NO: 131 and SEQ ID NO: 2.

A 3' UTR is region of an mRNA that is directly downstream (3') from the stop codon (the codon of an mRNA transcript that signals a termination of translation). A 3' UTR does not

encode a protein (is non-coding). Natural or wild type 3′ UTRs are known to have stretches of adenosines and uridines embedded in them. These AU rich signatures are particularly prevalent in genes with high rates of turnover. Based on their sequence features and functional properties, the AU rich elements (AREs) can be separated into three classes (Chen et al, 1995): Class I AREs contain several dispersed copies of an AUUUA motif within U-rich regions. C-Myc and MyoD contain class I AREs. Class II AREs possess two or more overlapping UUAUUUA(U/A)(U/A) (SEQ ID NO: 130) nonamers. Molecules containing this type of AREs include GM-CSF and TNF-a. Class III ARES are less well defined. These U rich regions do not contain an AUUUA motif. c-Jun and Myogenin are two well-studied examples of this class. Most proteins binding to the AREs are known to destabilize the messenger, whereas members of the ELAV family, most notably HuR, have been documented to increase the stability of mRNA. HuR binds to AREs of all the three classes. Engineering the HuR specific binding sites into the 3′ UTR of nucleic acid molecules will lead to HuR binding and thus, stabilization of the message in vivo.

Introduction, removal or modification of 3′ UTR AU rich elements (AREs) can be used to modulate the stability of nucleic acids (e.g., RNA) of the disclosure. When engineering specific nucleic acids, one or more copies of an ARE can be introduced to make nucleic acids of the disclosure less stable and thereby curtail translation and decrease production of the resultant protein. Likewise, AREs can be identified and removed or mutated to increase the intracellular stability and thus increase translation and production of the resultant protein. Transfection experiments can be conducted in relevant cell lines, using nucleic acids of the disclosure and protein production can be assayed at various time points post-transfection. For example, cells can be transfected with different ARE-engineering molecules and by using an ELISA kit to the relevant protein and assaying protein produced at 6 hour, 12 hour, 24 hour, 48 hour, and 7 days post-transfection.

3′ UTRs may be heterologous or synthetic. With respect to 3' UTRs, globin UTRs, including Xenopus β-globin UTRs and human β-globin UTRs are known in the art (8278063, 9012219, US20110086907). A modified β-globin construct with enhanced stability in some cell types by cloning two sequential human β-globin 3'UTRs head to tail has been developed and is well known in the art (US2012/0195936, WO2014/071963). Addtionally, a2-globin, a1-globin, UTRs and mutants thereof are also known in the art (WO2015101415, WO2015024667). Other 3′ UTRs described in the mRNA constructs in the non-patent literature include CYBA (Ferizi et al., 2015) and albumin (Thess et al., 2015). Other exemplary 3′ UTRs include that of bovine or human growth hormone (wild type or modified) (WO2013/185069, US20140206753,

WO2014152774), rabbit β globin and hepatitis B virus (HBV), α-globin 3′ UTR and Viral VEEV 3' UTR sequences are also known in the art. In some embodiments, the sequence UUUGAAUU (WO2014144196) is used. In some embodiments, 3′ UTRs of human and mouse ribosomal protein are used. Other examples include rps9 3'UTR (WO2015101414), FIG4 (WO2015101415), and human albumin 7 (WO2015101415).

In some embodiments, a 3' UTR of the present disclosure comprises a sequence selected from SEQ ID NO: 132 and SEQ ID NO: 4.

Those of ordinary skill in the art will understand that 5'UTRs that are heterologous or synthetic may be used with any desired 3' UTR sequence. For example, a heterologous 5'UTR may be used with a synthetic 3'UTR with a heterologous 3" UTR.

Non-UTR sequences may also be used as regions or subregions within a nucleic acid. For example, introns or portions of introns sequences may be incorporated into regions of nucleic acid of the disclosure. Incorporation of intronic sequences may increase protein production as well as nucleic acid levels.

Combinations of features may be included in flanking regions and may be contained within other features. For example, the ORF may be flanked by a 5' UTR which may contain a strong Kozak translational initiation signal and/or a 3' UTR which may include an oligo(dT) sequence for templated addition of a poly-A tail. 5′ UTR may comprise a first polynucleotide fragment and a second polynucleotide fragment from the same and/or different genes such as the 5′ UTRs described in US Patent Application Publication No.20100293625 and PCT/US2014/069155, herein incorporated by reference in its entirety.

It should be understood that any UTR from any gene may be incorporated into the regions of a nucleic acid. Furthermore, multiple wild-type UTRs of any known gene may be utilized. It is also within the scope of the present disclosure to provide artificial UTRs which are not variants of wild type regions. These UTRs or portions thereof may be placed in the same orientation as in the transcript from which they were selected or may be altered in orientation or location. Hence a 5′ or 3′ UTR may be inverted, shortened, lengthened, made with one or more other 5′ UTRs or 3′ UTRs. As used herein, the term "altered" as it relates to a UTR sequence, means that the UTR has been changed in some way in relation to a reference sequence. For example, a 3′ UTR or 5′ UTR may be altered relative to a wild-type or native UTR by the change in orientation or location as taught above or may be altered by the inclusion of additional nucleotides, deletion of nucleotides, swapping or transposition of nucleotides. Any of these changes producing an "altered" UTR (whether 3′ or 5′) comprise a variant UTR.

In some embodiments, a double, triple or quadruple UTR such as a 5′ UTR or 3′ UTR may be used. As used herein, a "double" UTR is one in which two copies of the same UTR are encoded either in series or substantially in series. For example, a double beta-globin 3′ UTR may be used as described in US Patent publication 20100129877, the contents of which are incorporated herein by reference in its entirety.

It is also within the scope of the present disclosure to have patterned UTRs. As used herein "patterned UTRs" are those UTRs which reflect a repeating or alternating pattern, such as ABABAB or AABBAABBAABB or ABCABCABC or variants thereof repeated once, twice, or more than 3 times. In these patterns, each letter, A, B, or C represent a different UTR at the nucleotide level.

In some embodiments, flanking regions are selected from a family of transcripts whose proteins share a common function, structure, feature or property. For example, polypeptides of interest may belong to a family of proteins which are expressed in a particular cell, tissue or at some time during development. The UTRs from any of these genes may be swapped for any other UTR of the same or different family of proteins to create a new polynucleotide. As used herein, a "family of proteins" is used in the broadest sense to refer to a group of two or more polypeptides of interest which share at least one function, structure, feature, localization, origin, or expression pattern.

The untranslated region may also include translation enhancer elements (TEE). As a non-limiting example, the TEE may include those described in US Application No.20090226470, herein incorporated by reference in its entirety, and those known in the art.

In vitro Transcription of RNA

cDNA encoding the polynucleotides described herein may be transcribed using an in vitro transcription (IVT) system. In vitro transcription of RNA is known in the art and is described in International Publication WO 2014/152027, which is incorporated by reference herein in its entirety. In some embodiments, the RNA of the present disclosure is prepared in accordance with any one or more of the methods described in WO 2018/053209 and WO 2019/036682, each of which is incorporated by reference herein.

In some embodiments, the RNA transcript is generated using a non-amplified, linearized DNA template in an in vitro transcription reaction to generate the RNA transcript. In some embodiments, the template DNA is isolated DNA. In some embodiments, the template DNA is cDNA. In some embodiments, the cDNA is formed by reverse transcription of an mRNA, for example, but not limited to coronavirus mRNA. In some embodiments, cells, e.g., bacterial cells,

e.g., E. coli, e.g., DH-1 cells are transfected with the plasmid DNA template. In some embodiments, the transfected cells are cultured to replicate the plasmid DNA which is then isolated and purified. In some embodiments, the DNA template includes an RNA polymerase promoter, e.g., a T7 promoter located 5 ' to and operably linked to the gene of interest.

In some embodiments, an in vitro transcription template encodes a 5′ untranslated (UTR) region, contains an open reading frame, and encodes a 3′ UTR and a poly(A) tail. The particular nucleic acid sequence composition and length of an in vitro transcription template will depend on the mRNA encoded by the template.

A "5′ untranslated region" (UTR) refers to a region of an mRNA that is directly upstream (i.e., 5′) from the start codon (i.e., the first codon of an mRNA transcript translated by a ribosome) that does not encode a polypeptide. When RNA transcripts are being generated, the 5' UTR may comprise a promoter sequence. Such promoter sequences are known in the art. It should be understood that such promoter sequences will not be present in a vaccine of the disclosure.

A "3′ untranslated region" (UTR) refers to a region of an mRNA that is directly downstream (i.e., 3′) from the stop codon (i.e., the codon of an mRNA transcript that signals a termination of translation) that does not encode a polypeptide.

An "open reading frame" is a continuous stretch of DNA beginning with a start codon (e.g., methionine (ATG)), and ending with a stop codon (e.g., TAA, TAG or TGA) and encodes a polypeptide.

A "poly(A) tail" is a region of mRNA that is downstream, e.g., directly downstream (i.e., 3′), from the 3′ UTR that contains multiple, consecutive adenosine monophosphates. A poly(A) tail may contain 10 to 300 adenosine monophosphates. For example, a poly(A) tail may contain 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290 or 300 adenosine monophosphates. In some embodiments, a poly(A) tail contains 50 to 250 adenosine monophosphates. In a relevant biological setting (e.g., in cells, in vivo) the poly(A) tail functions to protect mRNA from enzymatic degradation, e.g., in the cytoplasm, and aids in transcription termination, and/or export of the mRNA from the nucleus and translation.

In some embodiments, a nucleic acid includes 200 to 3,000 nucleotides. For example, a nucleic acid may include 200 to 500, 200 to 1000, 200 to 1500, 200 to 3000, 500 to 1000, 500 to 1500, 500 to 2000, 500 to 3000, 1000 to 1500, 1000 to 2000, 1000 to 3000, 1500 to 3000, or 2000 to 3000 nucleotides).

An in vitro transcription system typically comprises a transcription buffer, nucleotide triphosphates (NTPs), an RNase inhibitor and a polymerase.

The NTPs may be manufactured in house, may be selected from a supplier, or may be synthesized as described herein. The NTPs may be selected from, but are not limited to, those described herein including natural and unnatural (modified) NTPs.

Any number of RNA polymerases or variants may be used in the method of the present disclosure. The polymerase may be selected from, but is not limited to, a phage RNA polymerase, e.g., a T7 RNA polymerase, a T3 RNA polymerase, a SP6 RNA polymerase, and/or mutant polymerases such as, but not limited to, polymerases able to incorporate modified nucleic acids and/or modified nucleotides, including chemically modified nucleic acids and/or nucleotides. Some embodiments exclude the use of DNase.

In some embodiments, the RNA transcript is capped via enzymatic capping. In some embodiments, the RNA comprises 5' terminal cap, for example, 7mG(5')ppp(5')NlmpNp.

Chemical Synthesis

Solid-phase chemical synthesis. Nucleic acids the present disclosure may be manufactured in whole or in part using solid phase techniques. Solid-phase chemical synthesis of nucleic acids is an automated method wherein molecules are immobilized on a solid support and synthesized step by step in a reactant solution. Solid-phase synthesis is useful in site-specific introduction of chemical modifications in the nucleic acid sequences.

Liquid Phase Chemical Synthesis. The synthesis of nucleic acids of the present disclosure by the sequential addition of monomer building blocks may be carried out in a liquid phase.

Combination of Synthetic Methods. The synthetic methods discussed above each has its own advantages and limitations. Attempts have been conducted to combine these methods to overcome the limitations. Such combinations of methods are within the scope of the present disclosure. The use of solid-phase or liquid-phase chemical synthesis in combination with enzymatic ligation provides an efficient way to generate long chain nucleic acids that cannot be obtained by chemical synthesis alone.

Ligation of Nucleic Acid Regions or Subregions

Assembling nucleic acids by a ligase may also be used. DNA or RNA ligases promote intermolecular ligation of the 5' and 3' ends of polynucleotide chains through the formation of a phosphodiester bond. Nucleic acids such as chimeric polynucleotides and/or circular nucleic

acids may be prepared by ligation of one or more regions or subregions. DNA fragments can be joined by a ligase catalyzed reaction to create recombinant DNA with different functions. Two oligodeoxynucleotides, one with a 5' phosphoryl group and another with a free 3' hydroxyl group, serve as substrates for a DNA ligase.

Purification

Purification of the nucleic acids described herein may include, but is not limited to, nucleic acid clean-up, quality assurance and quality control. Clean-up may be performed by methods known in the arts such as, but not limited to, AGENCOURT® beads (Beckman Coulter Genomics, Danvers, MA), poly-T beads, LNATM oligo-T capture probes (EXIQON® Inc, Vedbaek, Denmark) or HPLC based purification methods such as, but not limited to, strong anion exchange HPLC, weak anion exchange HPLC, reverse phase HPLC (RP-HPLC), and hydrophobic interaction HPLC (HIC-HPLC). The term "purified" when used in relation to a nucleic acid such as a "purified nucleic acid" refers to one that is separated from at least one contaminant. A "contaminant" is any substance that makes another unfit, impure or inferior. Thus, a purified nucleic acid (e.g., DNA and RNA) is present in a form or setting different from that in which it is found in nature, or a form or setting different from that which existed prior to subjecting it to a treatment or purification method.

A quality assurance and/or quality control check may be conducted using methods such as, but not limited to, gel electrophoresis, UV absorbance, or analytical HPLC.

In some embodiments, the nucleic acids may be sequenced by methods including, but not limited to reverse-transcriptase-PCR.

Quantification

In some embodiments, the nucleic acids of the present disclosure may be quantified in exosomes or when derived from one or more bodily fluid. Bodily fluids include peripheral blood, serum, plasma, ascites, urine, cerebrospinal fluid (CSF), sputum, saliva, bone marrow, synovial fluid, aqueous humor, amniotic fluid, cerumen, breast milk, broncheoalveolar lavage fluid, semen, prostatic fluid, cowper's fluid or pre-ejaculatory fluid, sweat, fecal matter, hair, tears, cyst fluid, pleural and peritoneal fluid, pericardial fluid, lymph, chyme, chyle, bile, interstitial fluid, menses, pus, sebum, vomit, vaginal secretions, mucosal secretion, stool water, pancreatic juice, lavage fluids from sinus cavities, bronchopulmonary aspirates, blastocyl cavity fluid, and umbilical cord blood. Alternatively, exosomes may be retrieved from an organ selected from the

group consisting of lung, heart, pancreas, stomach, intestine, bladder, kidney, ovary, testis, skin, colon, breast, prostate, brain, esophagus, liver, and placenta.

Assays may be performed using construct specific probes, cytometry, qRT-PCR, real-time PCR, PCR, flow cytometry, electrophoresis, mass spectrometry, or combinations thereof while the exosomes may be isolated using immunohistochemical methods such as enzyme linked immunosorbent assay (ELISA) methods. Exosomes may also be isolated by size exclusion chromatography, density gradient centrifugation, differential centrifugation, nanomembrane ultrafiltration, immunoabsorbent capture, affinity purification, microfluidic separation, or combinations thereof.

These methods afford the investigator the ability to monitor, in real time, the level of nucleic acids remaining or delivered. This is possible because the nucleic acids of the present disclosure, in some embodiments, differ from the endogenous forms due to the structural or chemical modifications.

In some embodiments, the nucleic acid may be quantified using methods such as, but not limited to, ultraviolet visible spectroscopy (UV/Vis). A non-limiting example of a UV/Vis spectrometer is a NANODROP® spectrometer (ThermoFisher, Waltham, MA). The quantified nucleic acid may be analyzed in order to determine if the nucleic acid may be of proper size, check that no degradation of the nucleic acid has occurred. Degradation of the nucleic acid may be checked by methods such as, but not limited to, agarose gel electrophoresis, HPLC based purification methods such as, but not limited to, strong anion exchange HPLC, weak anion exchange HPLC, reverse phase HPLC (RP-HPLC), and hydrophobic interaction HPLC (HIC-HPLC), liquid chromatography-mass spectrometry (LCMS), capillary electrophoresis (CE) and capillary gel electrophoresis (CGE).

Lipid Nanoparticles (LNPs)

In some embodiments, the mRNA of the disclosure is formulated in a lipid nanoparticle (LNP). Lipid nanoparticles typically comprise ionizable cationic lipid, non-cationic lipid, sterol and PEG lipid components along with the nucleic acid cargo of interest. The lipid nanoparticles of the disclosure can be generated using components, compositions, and methods as are generally known in the art, see for example PCT/US2016/052352; PCT/US2016/068300; PCT/US2017/037551; PCT/US2015/027400; PCT/US2016/047406; PCT/US2016000129; PCT/US2016/014280; PCT/US2016/014280; PCT/US2017/038426; PCT/US2014/027077; PCT/US2014/055394; PCT/US2016/52117; PCT/US2012/069610; PCT/US2017/027492;

PCT/US2016/059575 and PCT/US2016/069491 all of which are incorporated by reference herein in their entirety.

Vaccines of the present disclosure are typically formulated in lipid nanoparticle. In some embodiments, the lipid nanoparticle comprises at least one ionizable cationic lipid, at least one non-cationic lipid, at least one sterol, and/or at least one polyethylene glycol (PEG)-modified lipid.

In some embodiments, the lipid nanoparticle comprises 40-50 mol% ionizable lipid, optionally 45-50 mol%, for example, 45-46 mol%, 46-47 mol%, 47-48 mol%, 48-49 mol%, or 49-50 mol% for example about 45 mol%, 45.5 mol%, 46 mol%, 46.5 mol%, 47 mol%, 47.5 mol%, 48 mol%, 48.5 mol%, 49 mol%, or 49.5 mol%.

In some embodiments, the lipid nanoparticle comprises 30-45 mol% sterol, optionally 35-40 mol%, for example, 30-31 mol%, 31-32 mol%, 32-33 mol%, 33-34 mol%, 35-35 mol%, 35-36 mol%, 36-37 mol%, 38-38 mol%, 38-39 mol%, or 39-40 mol%.

In some embodiments, the lipid nanoparticle comprises 5-15 mol% helper lipid, optionally 10-12 mol%, for example, 5-6 mol%, 6-7 mol%, 7-8 mol%, 8-9 mol%, 9-10 mol%, 10-11 mol%, 11-12 mol%, 12-13 mol%, 13-14 mol%, or 14-15 mol%.

In some embodiments, the lipid nanoparticle comprises 1-5% PEG lipid, optionally 1-3 mol%, for example 1.5 to 2.5 mol%, 1-2 mol%, 2-3 mol%, 3-4 mol%, or 4-5 mol%.

In some embodiments, the lipid nanoparticle comprises 20-60 mol% ionizable cationic lipid. For example, the lipid nanoparticle may comprise 20-50 mol%, 20-40 mol%, 20-30 mol%, 30-60 mol%, 30-50 mol%, 30-40 mol%, 40-60 mol%, 40-50 mol%, or 50-60 mol% ionizable cationic lipid. In some embodiments, the lipid nanoparticle comprises 20 mol%, 30 mol%, 40 mol%, 50 mol%, or 60 mol% ionizable cationic lipid. In some embodiments, the lipid nanoparticle comprises 35 mol%, 36 mol%, 37 mol%, 38 mol%, 39 mol%, 40 mol%, 41 mol%, 42 mol%, 43 mol%, 44 mol%, 45 mol%, 46 mol%, 47 mol%, 48 mol%, 49 mol%, 50 mol%, 51 mol%, 52 mol%, 53 mol%, 54 mol%, or 55 mol% ionizable cationic lipid.

In some embodiments, the lipid nanoparticle comprises 5-25 mol% non-cationic lipid. For example, the lipid nanoparticle may comprise 5-20 mol%, 5-15 mol%, 5-10 mol%, 10-25 mol%, 10-20 mol%, 10-25 mol%, 15-25 mol%, 15-20 mol%, or 20-25 mol% non-cationic lipid. In some embodiments, the lipid nanoparticle comprises 5 mol%, 10 mol%, 15 mol%, 20 mol%, or 25 mol% non-cationic lipid.

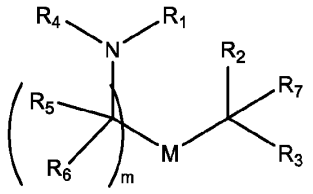
In some embodiments, the lipid nanoparticle comprises 25-55 mol% sterol. For example, the lipid nanoparticle may comprise 25-50 mol%, 25-45 mol%, 25-40 mol%, 25-35 mol%, 25-30 mol%, 30-55 mol%, 30-50 mol%, 30-45 mol%, 30-40 mol%, 30-35 mol%, 35-55 mol%, 35-50

mol%, 35-45 mol%, 35-40 mol%, 40-55 mol%, 40-50 mol%, 40-45 mol%, 45-55 mol%, 45-50 mol%, or 50-55 mol% sterol. In some embodiments, the lipid nanoparticle comprises 25 mol%, 30 mol%, 35 mol%, 40 mol%, 45 mol%, 50 mol%, or 55 mol% sterol.

In some embodiments, the lipid nanoparticle comprises 0.5-15 mol% PEG-modified lipid. For example, the lipid nanoparticle may comprise 0.5-10 mol%, 0.5-5 mol%, 1-15 mol%, 1-10 mol%, 1-5 mol%, 2-15 mol%, 2-10 mol%, 2-5 mol%, 5-15 mol%, 5-10 mol%, or 10-15 mol%. In some embodiments, the lipid nanoparticle comprises 0.5 mol%, 1 mol%, 2 mol%, 3 mol%, 4 mol%, 5 mol%, 6 mol%, 7 mol%, 8 mol%, 9 mol%, 10 mol%, 11 mol%, 12 mol%, 13 mol%, 14 mol%, or 15 mol% PEG-modified lipid.

In some embodiments, the lipid nanoparticle comprises 20-60 mol% ionizable cationic lipid, 5-25 mol% non-cationic lipid, 25-55 mol% sterol, and 0.5-15 mol% PEG-modified lipid.

In some embodiments, an ionizable cationic lipid of the disclosure comprises a compound of Formula (I):



(I),

or a salt or isomer thereof, wherein:

R_1 is selected from the group consisting of C_{5-30} alkyl, C_{5-20} alkenyl, -R*YR", -YR", and -R"M'R';

R_2 and R_3 are independently selected from the group consisting of H, C_{1-14} alkyl, C_{2-14} alkenyl, -R*YR", -YR", and -R*OR", or R_2 and R_3, together with the atom to which they are attached, form a heterocycle or carbocycle;

R_4 is selected from the group consisting of a C_{3-6} carbocycle, $-(CH_2)_nQ$, $-(CH_2)_nCHQR$, -CHQR, $-CQ(R)_2$, and unsubstituted C_{1-6} alkyl, where Q is selected from a carbocycle, heterocycle, -OR, $-O(CH_2)_nN(R)_2$, -C(O)OR, -OC(O)R, $-CX_3$, $-CX_2H$, $-CXH_2$, -CN, $-N(R)_2$, $-C(O)N(R)_2$, $-N(R)C(O)R$, $-N(R)S(O)_2R$, $-N(R)C(O)N(R)_2$, $-N(R)C(S)N(R)_2$, $-N(R)R_8$, $-O(CH_2)_nOR$, $-N(R)C(=NR_9)N(R)_2$, $-N(R)C(=CHR_9)N(R)_2$, $-OC(O)N(R)_2$, -N(R)C(O)OR, -N(OR)C(O)R, $-N(OR)S(O)_2R$, -N(OR)C(O)OR, $-N(OR)C(O)N(R)_2$, $-N(OR)C(S)N(R)_2$, $-N(OR)C(=NR_9)N(R)_2$, $-N(OR)C(=CHR_9)N(R)_2$, $-C(=NR_9)N(R)_2$, $-C(=NR_9)R$, -C(O)N(R)OR, and $-C(R)N(R)_2C(O)OR$, and each n is independently selected from 1, 2, 3, 4, and 5;

each R_5 is independently selected from the group consisting of C_{1-3} alkyl, C_{2-3} alkenyl, and H;

each R_6 is independently selected from the group consisting of C_{1-3} alkyl, C_{2-3} alkenyl, and H;

M and M' are independently selected from -C(O)O-, -OC(O)-, -C(O)N(R')-, -N(R')C(O)-, -C(O)-, -C(S)-, -C(S)S-, -SC(S)-, -CH(OH)-, -P(O)(OR')O-, -S(O)$_2$-, -S-S-, an aryl group, and a heteroaryl group;

R_7 is selected from the group consisting of C_{1-3} alkyl, C_{2-3} alkenyl, and H;

R_8 is selected from the group consisting of C_{3-6} carbocycle and heterocycle;

R_9 is selected from the group consisting of H, CN, NO_2, C_{1-6} alkyl, -OR, -S(O)$_2$R, -S(O)$_2$N(R)$_2$, C_{2-6} alkenyl, C_{3-6} carbocycle and heterocycle;

each R is independently selected from the group consisting of C_{1-3} alkyl, C_{2-3} alkenyl, and H;

each R' is independently selected from the group consisting of C_{1-18} alkyl, C_{2-18} alkenyl, -R*YR", -YR", and H;

each R" is independently selected from the group consisting of C_{3-14} alkyl and C_{3-14} alkenyl;

each R* is independently selected from the group consisting of C_{1-12} alkyl and C_{2-12} alkenyl;

each Y is independently a C_{3-6} carbocycle;

each X is independently selected from the group consisting of F, Cl, Br, and I; and

m is selected from 5, 6, 7, 8, 9, 10, 11, 12, and 13.

In some embodiments, a subset of compounds of Formula (I) includes those in which when R_4 is -(CH$_2$)$_n$Q, -(CH$_2$)$_n$CHQR, –CHQR, or -CQ(R)$_2$, then (i) Q is not -N(R)$_2$ when n is 1, 2, 3, 4 or 5, or (ii) Q is not 5, 6, or 7-membered heterocycloalkyl when n is 1 or 2.

In some embodiments, another subset of compounds of Formula (I) includes those in which

R_1 is selected from the group consisting of C_{5-30} alkyl, C_{5-20} alkenyl, -R*YR", -YR", and -R"M'R';

R_2 and R_3 are independently selected from the group consisting of H, C_{1-14} alkyl, C_{2-14} alkenyl, -R*YR", -YR", and -R*OR", or R_2 and R_3, together with the atom to which they are attached, form a heterocycle or carbocycle;

R_4 is selected from the group consisting of a C_{3-6} carbocycle, -(CH$_2$)$_n$Q, -(CH$_2$)$_n$CHQR, -CHQR, -CQ(R)$_2$, and unsubstituted C_{1-6} alkyl, where Q is selected from a C_{3-6} carbocycle, a 5- to 14-membered heteroaryl having one or more heteroatoms selected from N, O, and S, -OR, -O(CH$_2$)$_n$N(R)$_2$, -C(O)OR, -OC(O)R, -CX$_3$, -CX$_2$H, -CXH$_2$, -CN, -C(O)N(R)$_2$,

-N(R)C(O)R, -N(R)S(O)$_2$R, -N(R)C(O)N(R)$_2$, -N(R)C(S)N(R)$_2$, -CRN(R)$_2$C(O)OR, -N(R)R$_8$, -O(CH$_2$)$_n$OR, -N(R)C(=NR$_9$)N(R)$_2$, -N(R)C(=CHR$_9$)N(R)$_2$, -OC(O)N(R)$_2$, -N(R)C(O)OR, -N(OR)C(O)R, -N(OR)S(O)$_2$R, -N(OR)C(O)OR, -N(OR)C(O)N(R)$_2$, -N(OR)C(S)N(R)$_2$, -N(OR)C(=NR$_9$)N(R)$_2$, -N(OR)C(=CHR$_9$)N(R)$_2$, -C(=NR$_9$)N(R)$_2$, -C(=NR$_9$)R, -C(O)N(R)OR, and a 5- to 14-membered heterocycloalkyl having one or more heteroatoms selected from N, O, and S which is substituted with one or more substituents selected from oxo (=O), OH, amino, mono- or di-alkylamino, and C$_{1-3}$ alkyl, and each n is independently selected from 1, 2, 3, 4, and 5;

each R$_5$ is independently selected from the group consisting of C$_{1-3}$ alkyl, C$_{2-3}$ alkenyl, and H;

each R$_6$ is independently selected from the group consisting of C$_{1-3}$ alkyl, C$_{2-3}$ alkenyl, and H;

M and M' are independently selected from -C(O)O-, -OC(O)-, -C(O)N(R')-, -N(R')C(O)-, -C(O)-, -C(S)-, -C(S)S-, -SC(S)-, -CH(OH)-, -P(O)(OR')O-, -S(O)$_2$-, -S-S-, an aryl group, and a heteroaryl group;

R$_7$ is selected from the group consisting of C$_{1-3}$ alkyl, C$_{2-3}$ alkenyl, and H;

R$_8$ is selected from the group consisting of C$_{3-6}$ carbocycle and heterocycle;

R$_9$ is selected from the group consisting of H, CN, NO$_2$, C$_{1-6}$ alkyl, -OR, -S(O)$_2$R, -S(O)$_2$N(R)$_2$, C$_{2-6}$ alkenyl, C$_{3-6}$ carbocycle and heterocycle;

each R is independently selected from the group consisting of C$_{1-3}$ alkyl, C$_{2-3}$ alkenyl, and H;

each R' is independently selected from the group consisting of C$_{1-18}$ alkyl, C$_{2-18}$ alkenyl, -R*YR", -YR", and H;

each R" is independently selected from the group consisting of C$_{3-14}$ alkyl and C$_{3-14}$ alkenyl;

each R* is independently selected from the group consisting of C$_{1-12}$ alkyl and C$_{2-12}$ alkenyl;

each Y is independently a C$_{3-6}$ carbocycle;

each X is independently selected from the group consisting of F, Cl, Br, and I; and

m is selected from 5, 6, 7, 8, 9, 10, 11, 12, and 13,

or salts or isomers thereof.

In some embodiments, another subset of compounds of Formula (I) includes those in which

R_1 is selected from the group consisting of C_{5-30} alkyl, C_{5-20} alkenyl, -R*YR", -YR", and -R"M'R';

R_2 and R_3 are independently selected from the group consisting of H, C_{1-14} alkyl, C_{2-14} alkenyl, -R*YR", -YR", and -R*OR", or R_2 and R_3, together with the atom to which they are attached, form a heterocycle or carbocycle;

R_4 is selected from the group consisting of a C_{3-6} carbocycle, $-(CH_2)_nQ$, $-(CH_2)_nCHQR$, -CHQR, $-CQ(R)_2$, and unsubstituted C_{1-6} alkyl, where Q is selected from a C_{3-6} carbocycle, a 5- to 14-membered heterocycle having one or more heteroatoms selected from N, O, and S, -OR, $-O(CH_2)_nN(R)_2$, -C(O)OR, -OC(O)R, $-CX_3$, $-CX_2H$, $-CXH_2$, -CN, $-C(O)N(R)_2$, -N(R)C(O)R, $-N(R)S(O)_2R$, $-N(R)C(O)N(R)_2$, $-N(R)C(S)N(R)_2$, $-CRN(R)_2C(O)OR$, $-N(R)R_8$, $-O(CH_2)_nOR$, $-N(R)C(=NR_9)N(R)_2$, $-N(R)C(=CHR_9)N(R)_2$, $-OC(O)N(R)_2$, -N(R)C(O)OR, -N(OR)C(O)R, $-N(OR)S(O)_2R$, -N(OR)C(O)OR, $-N(OR)C(O)N(R)_2$, $-N(OR)C(S)N(R)_2$, $-N(OR)C(=NR_9)N(R)_2$, $-N(OR)C(=CHR_9)N(R)_2$, $-C(=NR_9)R$, -C(O)N(R)OR, and $-C(=NR_9)N(R)_2$, and each n is independently selected from 1, 2, 3, 4, and 5; and when Q is a 5- to 14-membered heterocycle and (i) R_4 is $-(CH_2)_nQ$ in which n is 1 or 2, or (ii) R_4 is $-(CH_2)_nCHQR$ in which n is 1, or (iii) R_4 is -CHQR, and $-CQ(R)_2$, then Q is either a 5- to 14-membered heteroaryl or 8- to 14-membered heterocycloalkyl;

each R_5 is independently selected from the group consisting of C_{1-3} alkyl, C_{2-3} alkenyl, and H;

each R_6 is independently selected from the group consisting of C_{1-3} alkyl, C_{2-3} alkenyl, and H;

M and M' are independently selected from -C(O)O-, -OC(O)-, -C(O)N(R')-, -N(R')C(O)-, -C(O)-, -C(S)-, -C(S)S-, -SC(S)-, -CH(OH)-, -P(O)(OR')O-, $-S(O)_2-$, -S-S-, an aryl group, and a heteroaryl group;

R_7 is selected from the group consisting of C_{1-3} alkyl, C_{2-3} alkenyl, and H;

R_8 is selected from the group consisting of C_{3-6} carbocycle and heterocycle;

R_9 is selected from the group consisting of H, CN, NO_2, C_{1-6} alkyl, -OR, $-S(O)_2R$, $-S(O)_2N(R)_2$, C_{2-6} alkenyl, C_{3-6} carbocycle and heterocycle;

each R is independently selected from the group consisting of C_{1-3} alkyl, C_{2-3} alkenyl, and H;

each R' is independently selected from the group consisting of C_{1-18} alkyl, C_{2-18} alkenyl, -R*YR", -YR", and H;

each R" is independently selected from the group consisting of C_{3-14} alkyl and C_{3-14} alkenyl;

each R* is independently selected from the group consisting of C_{1-12} alkyl and C_{2-12} alkenyl;

each Y is independently a C_{3-6} carbocycle;

each X is independently selected from the group consisting of F, Cl, Br, and I; and

m is selected from 5, 6, 7, 8, 9, 10, 11, 12, and 13,

or salts or isomers thereof.

In some embodiments, another subset of compounds of Formula (I) includes those in which

R_1 is selected from the group consisting of C_{5-30} alkyl, C_{5-20} alkenyl, -R*YR", -YR", and -R"M'R';

R_2 and R_3 are independently selected from the group consisting of H, C_{1-14} alkyl, C_{2-14} alkenyl, -R*YR", -YR", and -R*OR", or R_2 and R_3, together with the atom to which they are attached, form a heterocycle or carbocycle;

R_4 is selected from the group consisting of a C_{3-6} carbocycle, $-(CH_2)_nQ$, $-(CH_2)_nCHQR$, -CHQR, $-CQ(R)_2$, and unsubstituted C_{1-6} alkyl, where Q is selected from a C_{3-6} carbocycle, a 5- to 14-membered heteroaryl having one or more heteroatoms selected from N, O, and S, -OR, $-O(CH_2)_nN(R)_2$, -C(O)OR, -OC(O)R, $-CX_3$, $-CX_2H$, $-CXH_2$, -CN, $-C(O)N(R)_2$, $-N(R)C(O)R$, $-N(R)S(O)_2R$, $-N(R)C(O)N(R)_2$, $-N(R)C(S)N(R)_2$, $-CRN(R)_2C(O)OR$, $-N(R)R_8$, $-O(CH_2)_nOR$, $-N(R)C(=NR_9)N(R)_2$, $-N(R)C(=CHR_9)N(R)_2$, $-OC(O)N(R)_2$, -N(R)C(O)OR, $-N(OR)C(O)R$, $-N(OR)S(O)_2R$, -N(OR)C(O)OR, $-N(OR)C(O)N(R)_2$, $-N(OR)C(S)N(R)_2$, $-N(OR)C(=NR_9)N(R)_2$, $-N(OR)C(=CHR_9)N(R)_2$, $-C(=NR_9)R$, -C(O)N(R)OR, and $-C(=NR_9)N(R)_2$, and each n is independently selected from 1, 2, 3, 4, and 5;

each R_5 is independently selected from the group consisting of C_{1-3} alkyl, C_{2-3} alkenyl, and H;

each R_6 is independently selected from the group consisting of C_{1-3} alkyl, C_{2-3} alkenyl, and H;

M and M' are independently selected from -C(O)O-, -OC(O)-, -C(O)N(R')-, -N(R')C(O)-, -C(O)-, -C(S)-, -C(S)S-, -SC(S)-, -CH(OH)-, -P(O)(OR')O-, $-S(O)_2-$, -S-S-, an aryl group, and a heteroaryl group;

R_7 is selected from the group consisting of C_{1-3} alkyl, C_{2-3} alkenyl, and H;

R_8 is selected from the group consisting of C_{3-6} carbocycle and heterocycle;

R_9 is selected from the group consisting of H, CN, NO_2, C_{1-6} alkyl, -OR, $-S(O)_2R$, $-S(O)_2N(R)_2$, C_{2-6} alkenyl, C_{3-6} carbocycle and heterocycle;

each R is independently selected from the group consisting of C_{1-3} alkyl, C_{2-3} alkenyl, and H;

each R' is independently selected from the group consisting of C_{1-18} alkyl, C_{2-18} alkenyl, -R*YR", -YR", and H;

each R" is independently selected from the group consisting of C_{3-14} alkyl and C_{3-14} alkenyl;

each R* is independently selected from the group consisting of C_{1-12} alkyl and C_{2-12} alkenyl;

each Y is independently a C_{3-6} carbocycle;

each X is independently selected from the group consisting of F, Cl, Br, and I; and

m is selected from 5, 6, 7, 8, 9, 10, 11, 12, and 13,

or salts or isomers thereof.

In some embodiments, another subset of compounds of Formula (I) includes those in which

R_1 is selected from the group consisting of C_{5-30} alkyl, C_{5-20} alkenyl, -R*YR", -YR", and -R"M'R';

R_2 and R_3 are independently selected from the group consisting of H, C_{2-14} alkyl, C_{2-14} alkenyl, -R*YR", -YR", and -R*OR", or R_2 and R_3, together with the atom to which they are attached, form a heterocycle or carbocycle;

R_4 is $-(CH_2)_nQ$ or $-(CH_2)_nCHQR$, where Q is $-N(R)_2$, and n is selected from 3, 4, and 5;

each R_5 is independently selected from the group consisting of C_{1-3} alkyl, C_{2-3} alkenyl, and H;

each R_6 is independently selected from the group consisting of C_{1-3} alkyl, C_{2-3} alkenyl, and H;

M and M' are independently selected from -C(O)O-, -OC(O)-, -C(O)N(R')-, -N(R')C(O)-, -C(O)-, -C(S)-, -C(S)S-, -SC(S)-, -CH(OH)-, -P(O)(OR')O-, $-S(O)_2-$, -S-S-, an aryl group, and a heteroaryl group;

R_7 is selected from the group consisting of C_{1-3} alkyl, C_{2-3} alkenyl, and H;

each R is independently selected from the group consisting of C_{1-3} alkyl, C_{2-3} alkenyl, and H;

each R' is independently selected from the group consisting of C_{1-18} alkyl, C_{2-18} alkenyl, -R*YR", -YR", and H;

each R" is independently selected from the group consisting of C_{3-14} alkyl and C_{3-14} alkenyl;

each R* is independently selected from the group consisting of C_{1-12} alkyl and C_{1-12} alkenyl;

each Y is independently a C_{3-6} carbocycle;

each X is independently selected from the group consisting of F, Cl, Br, and I; and

m is selected from 5, 6, 7, 8, 9, 10, 11, 12, and 13,

or salts or isomers thereof.

In some embodiments, another subset of compounds of Formula (I) includes those in which

R_1 is selected from the group consisting of C_{5-30} alkyl, C_{5-20} alkenyl, -R*YR", -YR", and -R"M'R';

R_2 and R_3 are independently selected from the group consisting of C_{1-14} alkyl, C_{2-14} alkenyl, -R*YR", -YR", and -R*OR", or R_2 and R_3, together with the atom to which they are attached, form a heterocycle or carbocycle;

R_4 is selected from the group consisting of $-(CH_2)_nQ$, $-(CH_2)_nCHQR$, -CHQR, and $-CQ(R)_2$, where Q is $-N(R)_2$, and n is selected from 1, 2, 3, 4, and 5;

each R_5 is independently selected from the group consisting of C_{1-3} alkyl, C_{2-3} alkenyl, and H;

each R_6 is independently selected from the group consisting of C_{1-3} alkyl, C_{2-3} alkenyl, and H;

M and M' are independently selected from -C(O)O-, -OC(O)-, -C(O)N(R')-, -N(R')C(O)-, -C(O)-, -C(S)-, -C(S)S-, -SC(S)-, -CH(OH)-, -P(O)(OR')O-, $-S(O)_2-$, -S-S-, an aryl group, and a heteroaryl group;

R_7 is selected from the group consisting of C_{1-3} alkyl, C_{2-3} alkenyl, and H;

each R is independently selected from the group consisting of C_{1-3} alkyl, C_{2-3} alkenyl, and H;

each R' is independently selected from the group consisting of C_{1-18} alkyl, C_{2-18} alkenyl, -R*YR", -YR", and H;

each R" is independently selected from the group consisting of C_{3-14} alkyl and C_{3-14} alkenyl;

each R* is independently selected from the group consisting of C_{1-12} alkyl and C_{1-12} alkenyl;

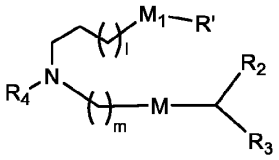
each Y is independently a C_{3-6} carbocycle;

each X is independently selected from the group consisting of F, Cl, Br, and I; and

m is selected from 5, 6, 7, 8, 9, 10, 11, 12, and 13,

or salts or isomers thereof.

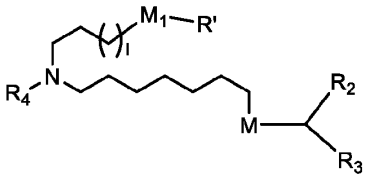
In some embodiments, a subset of compounds of Formula (I) includes those of Formula (IA):



(IA),

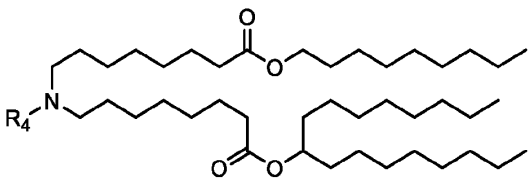
or a salt or isomer thereof, wherein l is selected from 1, 2, 3, 4, and 5; m is selected from 5, 6, 7, 8, and 9; M_1 is a bond or M'; R_4 is unsubstituted C_{1-3} alkyl, or $-(CH_2)_nQ$, in which Q is OH, $-NHC(S)N(R)_2$, $-NHC(O)N(R)_2$, $-N(R)C(O)R$, $-N(R)S(O)_2R$, $-N(R)R_8$, $-NHC(=NR_9)N(R)_2$, $-NHC(=CHR_9)N(R)_2$, $-OC(O)N(R)_2$, $-N(R)C(O)OR$, heteroaryl or heterocycloalkyl; M and M' are independently selected from -C(O)O-, -OC(O)-, -C(O)N(R')-, -P(O)(OR')O-, -S-S-, an aryl group, and a heteroaryl group; and R_2 and R_3 are independently selected from the group consisting of H, C_{1-14} alkyl, and C_{2-14} alkenyl.

In some embodiments, a subset of compounds of Formula (I) includes those of Formula (II):

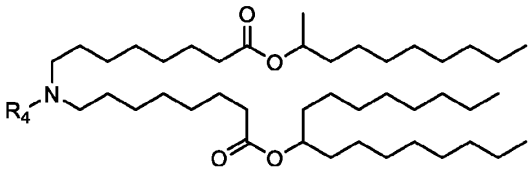


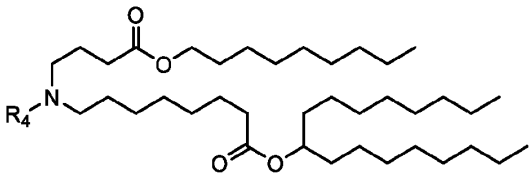
(II) or a salt or isomer thereof, wherein l is selected from 1, 2, 3, 4, and 5; M_1 is a bond or M'; R_4 is unsubstituted C_{1-3} alkyl, or $-(CH_2)_nQ$, in which n is 2, 3, or 4, and Q is OH, $-NHC(S)N(R)_2$, $-NHC(O)N(R)_2$, $-N(R)C(O)R$, $-N(R)S(O)_2R$, $-N(R)R_8$, $-NHC(=NR_9)N(R)_2$, $-NHC(=CHR_9)N(R)_2$, $-OC(O)N(R)_2$, $-N(R)C(O)OR$, heteroaryl or heterocycloalkyl; M and M' are independently selected from -C(O)O-, -OC(O)-, -C(O)N(R')-, -P(O)(OR')O-, -S-S-, an aryl group, and a heteroaryl group; and R_2 and R_3 are independently selected from the group consisting of H, C_{1-14} alkyl, and C_{2-14} alkenyl.

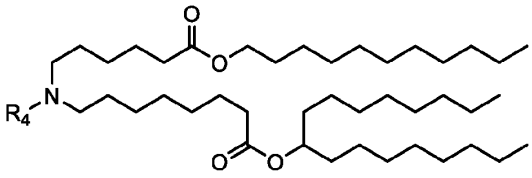
In some embodiments, a subset of compounds of Formula (I) includes those of Formula (IIa), (IIb), (IIc), or (IIe):



(IIa),

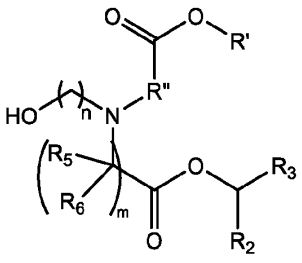

(IIb),


(IIc), or


(IIe),

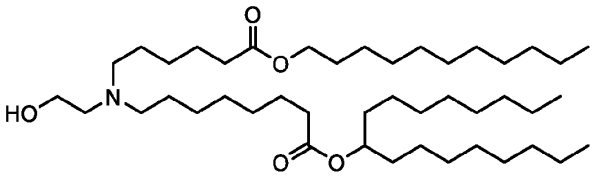
or a salt or isomer thereof, wherein R_4 is as described herein.

In some embodiments, a subset of compounds of Formula (I) includes those of Formula (IId):

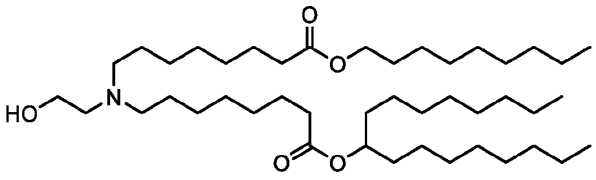

(IId),

or a salt or isomer thereof, wherein n is 2, 3, or 4; and m, R', R", and R_2 through R_6 are as described herein. For example, each of R_2 and R_3 may be independently selected from the group consisting of C_{5-14} alkyl and C_{5-14} alkenyl.

In some embodiments, an ionizable cationic lipid of the disclosure comprises a compound having structure:


(Compound I).

In some embodiments, an ionizable cationic lipid of the disclosure comprises a compound having structure:



(Compound II).

In some embodiments, a non-cationic lipid of the disclosure comprises 1,2-distearoyl-sn-glycero-3-phosphocholine (DSPC), 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine (DOPE), 1,2-dilinoleoyl-sn-glycero-3-phosphocholine (DLPC), 1,2-dimyristoyl-sn-gly cero-phosphocholine (DMPC), 1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPC), 1,2-dipalmitoyl-sn-glycero-3-phosphocholine (DPPC), 1,2-diundecanoyl-sn-glycero-phosphocholine (DUPC), 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC), 1,2-di-O-octadecenyl-sn-glycero-3-phosphocholine (18:0 Diether PC), 1-oleoyl-2 cholesterylhemisuccinoyl-sn-glycero-3-phosphocholine (OChemsPC), 1-hexadecyl-sn-glycero-3-phosphocholine (C16 Lyso PC), 1,2-dilinolenoyl-sn-glycero-3-phosphocholine,1,2-diarachidonoyl-sn-glycero-3-phosphocholine, 1,2-didocosahexaenoyl-sn-glycero-3-phosphocholine, 1,2-diphytanoyl-sn-glycero-3-phosphoethanolamine (ME 16.0 PE), 1,2-distearoyl-sn-glycero-3-phosphoethanolamine, 1,2-dilinoleoyl-sn-glycero-3-phosphoethanolamine, 1,2-dilinolenoyl-sn-glycero-3-phosphoethanolamine, 1,2-diarachidonoyl-sn-glycero-3-phosphoethanolamine, 1,2-didocosahexaenoyl-sn-glycero-3-phosphoethanolamine, 1,2-dioleoyl-sn-glycero-3-phospho-rac-(1-glycerol) sodium salt (DOPG), sphingomyelin, and mixtures thereof.

In some embodiments, a PEG modified lipid of the disclosure comprises a PEG-modified phosphatidylethanolamine, a PEG-modified phosphatidic acid, a PEG-modified ceramide, a PEG-modified dialkylamine, a PEG-modified diacylglycerol, a PEG-modified dialkylglycerol, and mixtures thereof. In some embodiments, the PEG-modified lipid is DMG-PEG, PEG-c-DOMG (also referred to as PEG-DOMG), PEG-DSG and/or PEG-DPG.

In some embodiments, a sterol of the disclosure comprises cholesterol, fecosterol, sitosterol, ergosterol, campesterol, stigmasterol, brassicasterol, tomatidine, ursolic acid, alpha-tocopherol, and mixtures thereof.

In some embodiments, a LNP of the disclosure comprises an ionizable cationic lipid of Compound 1, wherein the non-cationic lipid is DSPC, the structural lipid that is cholesterol, and the PEG lipid is DMG-PEG.

In some embodiments, the lipid nanoparticle comprises 45 – 55 mole percent (mol%) ionizable cationic lipid. For example, lipid nanoparticle may comprise 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, or 55 mol% ionizable cationic lipid.

In some embodiments, the lipid nanoparticle comprises 5 – 15 mol%, 5 – 10 mol%, or 10 – 15 mol% DSPC. For example, the lipid nanoparticle may comprise 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or 15 mol% DSPC.

In some embodiments, the lipid nanoparticle comprises 35 – 40 mol% cholesterol. For example, the lipid nanoparticle may comprise 35, 35.5, 36, 36.5, 37, 37.5, 38, 38.5, 39, 39.5, or 40 mol% cholesterol.

In some embodiments, the lipid nanoparticle comprises 1 – 2 mol%, 1 – 3 mol%, 1 – 4 mol%, or 1 – 5 mol% DMG-PEG. For example, the lipid nanoparticle may comprise 1, 1.5, 2, 2.5, 3, or 3.5 mol% DMG-PEG.

In some embodiments, the lipid nanoparticle comprises 50 mol% ionizable cationic lipid, 10 mol% DSPC, 38.5 mol% cholesterol, and 1.5 mol% DMG-PEG.

In some embodiments, the lipid nanoparticle comprises 49 mol% ionizable cationic lipid, 10 mol% DSPC, 38.5 mol% cholesterol, and 2.5 mol% DMG-PEG.

In some embodiments, the lipid nanoparticle comprises 49 mol% ionizable cationic lipid, 11 mol% DSPC, 38.5 mol% cholesterol, and 1.5 mol% DMG-PEG.

In some embodiments, the lipid nanoparticle comprises 48 mol% ionizable cationic lipid, 11 mol% DSPC, 38.5 mol% cholesterol, and 2.5 mol% DMG-PEG.

In some embodiments, an LNP of the disclosure comprises an N:P ratio of from about 2:1 to about 30:1.

In some embodiments, an LNP of the disclosure comprises an N:P ratio of about 6:1.

In some embodiments, an LNP of the disclosure comprises an N:P ratio of about 3:1.

In some embodiments, an LNP of the disclosure comprises a wt/wt ratio of the ionizable cationic lipid component to the RNA of from about 10:1 to about 100:1.

In some embodiments, an LNP of the disclosure comprises a wt/wt ratio of the ionizable cationic lipid component to the RNA of about 20:1.

In some embodiments, an LNP of the disclosure comprises a wt/wt ratio of the ionizable cationic lipid component to the RNA of about 10:1.

In some embodiments, an LNP of the disclosure has a mean diameter from about 50 nm to about 150 nm.

In some embodiments, an LNP of the disclosure has a mean diameter from about 70 nm to about 120 nm.

Multivalent Vaccines

The compositions, as provided herein, may include RNA or multiple RNAs encoding two or more antigens of the same or different species. In some embodiments, composition includes an mRNA or multiple mRNAs encoding two or more coronavirus antigens. In some embodiments, the RNA may encode 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, or more coronavirus antigens.

In some embodiments, two or more different mRNA encoding antigens may be formulated in the same lipid nanoparticle. In other embodiments, two or more different RNA encoding antigens may be formulated in separate lipid nanoparticles (each RNA formulated in a single lipid nanoparticle). The lipid nanoparticles may then be combined and administered as a single vaccine composition (e.g., comprising multiple RNA encoding multiple antigens) or may be administered separately.

Combination Vaccines

The compositions, as provided herein, may include an mRNA or multiple RNAs encoding two or more antigens of the same or different viral strains. Also provided herein are combination vaccines that include RNA encoding one or more coronavirus and one or more antigen(s) of a different organism. Thus, the vaccines of the present disclosure may be combination vaccines that target one or more antigens of the same strain/species, or one or more antigens of different strains/species, e.g., antigens which induce immunity to organisms which are found in the same geographic areas where the risk of coronavirus infection is high or organisms to which an individual is likely to be exposed to when exposed to a coronavirus.

Pharmaceutical Formulations

Provided herein are compositions (e.g., pharmaceutical compositions), methods, kits and reagents for prevention or treatment of coronavirus in humans and other mammals, for example. The compositions provided herein can be used as therapeutic or prophylactic agents. They may be used in medicine to prevent and/or treat a coronavirus infection.

In some embodiments, the coronavirus vaccine containing RNA as described herein can be administered to a subject (e.g., a mammalian subject, such as a human subject), and the mRNAs are translated in vivo to produce an antigenic polypeptide (antigen).

An "effective amount" of a composition (e.g., comprising RNA) is based, at least in part, on the target tissue, target cell type, means of administration, physical characteristics of the RNA (e.g., length, nucleotide composition, and/or extent of modified nucleosides), other components of the vaccine, and other determinants, such as age, body weight, height, sex and general health

of the subject. Typically, an effective amount of a composition provides an induced or boosted immune response as a function of antigen production in the cells of the subject. In some embodiments, an effective amount of the composition containing mRNA having at least one chemical modifications are more efficient than a composition containing a corresponding unmodified polynucleotide encoding the same antigen or a peptide antigen. Increased antigen production may be demonstrated by increased cell transfection (the percentage of cells transfected with the RNA vaccine), increased protein translation and/or expression from the polynucleotide, decreased nucleic acid degradation (as demonstrated, for example, by increased duration of protein translation from a modified polynucleotide), or altered antigen specific immune response of the host cell.

The term "pharmaceutical composition" refers to the combination of an active agent with a carrier, inert or active, making the composition especially suitable for diagnostic or therapeutic use in vivo or ex vivo. A "pharmaceutically acceptable carrier," after administered to or upon a subject, does not cause undesirable physiological effects. The carrier in the pharmaceutical composition must be "acceptable" also in the sense that it is compatible with the active ingredient and can be capable of stabilizing it. One or more solubilizing agents can be utilized as pharmaceutical carriers for delivery of an active agent. Examples of a pharmaceutically acceptable carrier include, but are not limited to, biocompatible vehicles, adjuvants, additives, and diluents to achieve a composition usable as a dosage form. Examples of other carriers include colloidal silicon oxide, magnesium stearate, cellulose, and sodium lauryl sulfate. Additional suitable pharmaceutical carriers and diluents, as well as pharmaceutical necessities for their use, are described in Remington's Pharmaceutical Sciences.

In some embodiments, the compositions (comprising polynucleotides and their encoded polypeptides) in accordance with the present disclosure may be used for treatment or prevention of a coronavirus infection. A composition may be administered prophylactically or therapeutically as part of an active immunization scheme to healthy individuals or early in infection during the incubation phase or during active infection after onset of symptoms. In some embodiments, the amount of RNA provided to a cell, a tissue or a subject may be an amount effective for immune prophylaxis.

A composition may be administered with other prophylactic or therapeutic compounds. As a non-limiting example, a prophylactic or therapeutic compound may be an adjuvant or a booster. As used herein, when referring to a prophylactic composition, such as a vaccine, the term “booster” refers to an extra administration of the prophylactic (vaccine) composition. A booster (or booster vaccine) may be given after an earlier administration of the prophylactic

composition. The time of administration between the initial administration of the prophylactic composition and the booster may be, but is not limited to, 1 minute, 2 minutes, 3 minutes, 4 minutes, 5 minutes, 6 minutes, 7 minutes, 8 minutes, 9 minutes, 10 minutes, 15 minutes, 20 minutes 35 minutes, 40 minutes, 45 minutes, 50 minutes, 55 minutes, 1 hour, 2 hours, 3 hours, 4 hours, 5 hours, 6 hours, 7 hours, 8 hours, 9 hours, 10 hours, 11 hours, 12 hours, 13 hours, 14 hours, 15 hours, 16 hours, 17 hours, 18 hours, 19 hours, 20 hours, 21 hours, 22 hours, 23 hours, 1 day, 36 hours, 2 days, 3 days, 4 days, 5 days, 6 days, 1 week, 10 days, 2 weeks, 3 weeks, 1 month, 2 months, 3 months, 4 months, 5 months, 6 months, 7 months, 8 months, 9 months, 10 months, 11 months, 1 year, 18 months, 2 years, 3 years, 4 years, 5 years, 6 years, 7 years, 8 years, 9 years, 10 years, 11 years, 12 years, 13 years, 14 years, 15 years, 16 years, 17 years, 18 years, 19 years, 20 years, 25 years, 30 years, 35 years, 40 years, 45 years, 50 years, 55 years, 60 years, 65 years, 70 years, 75 years, 80 years, 85 years, 90 years, 95 years or more than 99 years. In exemplary embodiments, the time of administration between the initial administration of the prophylactic composition and the booster may be, but is not limited to, 1 week, 2 weeks, 3 weeks, 1 month, 2 months, 3 months, 6 months or 1 year.

In some embodiments, a composition may be administered intramuscularly, intranasally or intradermally, similarly to the administration of inactivated vaccines known in the art.

A composition may be utilized in various settings depending on the prevalence of the infection or the degree or level of unmet medical need. As a non-limiting example, the RNA vaccines may be utilized to treat and/or prevent a variety of infectious disease. RNA vaccines have superior properties in that they produce much larger antibody titers, better neutralizing immunity, produce more durable immune responses, and/or produce responses earlier than commercially available vaccines.

Provided herein are pharmaceutical compositions including RNA and/or complexes optionally in combination with one or more pharmaceutically acceptable excipients.

The RNA may be formulated or administered alone or in conjunction with one or more other components. For example, a composition may comprise other components including, but not limited to, adjuvants.

In some embodiments, a composition does not include an adjuvant (they are adjuvant free).

An RNA may be formulated or administered in combination with one or more pharmaceutically-acceptable excipients. In some embodiments, vaccine compositions comprise at least one additional active substance, such as, for example, a therapeutically-active substance, a prophylactically-active substance, or a combination of both. Vaccine compositions may be

sterile, pyrogen-free or both sterile and pyrogen-free. General considerations in the formulation and/or manufacture of pharmaceutical agents, such as vaccine compositions, may be found, for example, in Remington: The Science and Practice of Pharmacy 21st ed., Lippincott Williams & Wilkins, 2005 (incorporated herein by reference in its entirety).

In some embodiments, a composition is administered to humans, human patients or subjects. For the purposes of the present disclosure, the phrase "active ingredient" generally refers to the RNA vaccines or the polynucleotides contained therein, for example, mRNA encoding antigens.

Formulations of the vaccine compositions described herein may be prepared by any method known or hereafter developed in the art of pharmacology. In general, such preparatory methods include the step of bringing the active ingredient (e.g., mRNA) into association with an excipient and/or one or more other accessory ingredients, and then, if necessary and/or desirable, dividing, shaping and/or packaging the product into a desired single- or multi-dose unit.

Relative amounts of the active ingredient, the pharmaceutically acceptable excipient, and/or any additional ingredients in a pharmaceutical composition in accordance with the disclosure will vary, depending upon the identity, size, and/or condition of the subject treated and further depending upon the route by which the composition is to be administered. By way of example, the composition may comprise between 0.1% and 100%, e.g., between 0.5 and 50%, between 1-30%, between 5-80%, at least 80% (w/w) active ingredient.

In some embodiments, an mRNA is formulated using one or more excipients to: (1) increase stability; (2) increase cell transfection; (3) permit the sustained or delayed release (e.g., from a depot formulation); (4) alter the biodistribution (e.g., target to specific tissues or cell types); (5) increase the translation of encoded protein in vivo; and/or (6) alter the release profile of encoded protein (antigen) in vivo. In addition to traditional excipients such as any and all solvents, dispersion media, diluents, or other liquid vehicles, dispersion or suspension aids, surface active agents, isotonic agents, thickening or emulsifying agents, preservatives, excipients can include, without limitation, lipidoids, liposomes, lipid nanoparticles, polymers, lipoplexes, core-shell nanoparticles, peptides, proteins, cells transfected with the RNA (e.g., for transplantation into a subject), hyaluronidase, nanoparticle mimics and combinations thereof.

Dosing/Administration

Provided herein are compositions (e.g., RNA vaccines), methods, kits and reagents for prevention and/or treatment of coronavirus infection in humans and other mammals. Immunizing compositions can be used as therapeutic or prophylactic agents. In some embodiments,

compositions are used to provide prophylactic protection from coronavirus infection. In some embodiments, compositions are used to treat a coronavirus infection. In some embodiments, embodiments, compositions are used in the priming of immune effector cells, for example, to activate peripheral blood mononuclear cells (PBMCs) ex vivo, which are then infused (re-infused) into a subject.

A subject may be any mammal, including non-human primate and human subjects. Typically, a subject is a human subject.

In some embodiments, a composition (e.g., RNA a vaccine) is administered to a subject (e.g., a mammalian subject, such as a human subject) in an effective amount to induce an antigen-specific immune response. The RNA encoding the coronavirus antigen is expressed and translated in vivo to produce the antigen, which then stimulates an immune response in the subject.

Prophylactic protection from a coronavirus can be achieved following administration of a composition of the present disclosure. Immunizing compositions can be administered once, twice, three times, four times or more but it is likely sufficient to administer the vaccine once (optionally followed by a single booster). It is possible, although less desirable, to administer a composition to an infected individual to achieve a therapeutic response. Dosing may need to be adjusted accordingly.

A method of eliciting an immune response in a subject against a coronavirus antigen (or multiple antigens) is provided in aspects of the present disclosure. In some embodiments, a method involves administering to the subject a composition comprising a mRNA having an open reading frame encoding a coronavirus antigen, thereby inducing in the subject an immune response specific to the coronavirus antigen, wherein anti-antigen antibody titer in the subject is increased following vaccination relative to anti-antigen antibody titer in a subject vaccinated with a prophylactically effective dose of a traditional vaccine against the antigen. An "anti-antigen antibody" is a serum antibody the binds specifically to the antigen.

A prophylactically effective dose is an effective dose that prevents infection with the virus at a clinically acceptable level. In some embodiments, the effective dose is a dose listed in a package insert for the vaccine. A traditional vaccine, as used herein, refers to a vaccine other than the mRNA vaccines of the present disclosure. For instance, a traditional vaccine includes, but is not limited, to live microorganism vaccines, killed microorganism vaccines, subunit vaccines, protein antigen vaccines, DNA vaccines, virus like particle (VLP) vaccines, etc. In exemplary embodiments, a traditional vaccine is a vaccine that has achieved regulatory approval

and/or is registered by a national drug regulatory body, for example the Food and Drug Administration (FDA) in the United States or the European Medicines Agency (EMA).

In some embodiments, the anti-antigen antibody titer in the subject is increased 1 log to 10 log following vaccination relative to anti-antigen antibody titer in a subject vaccinated with a prophylactically effective dose of a traditional vaccine against the coronavirus or an unvaccinated subject. In some embodiments, the anti-antigen antibody titer in the subject is increased 1 log, 2 log, 3 log, 4 log, 5 log, or 10 log following vaccination relative to anti-antigen antibody titer in a subject vaccinated with a prophylactically effective dose of a traditional vaccine against the coronavirus or an unvaccinated subject.

A method of eliciting an immune response in a subject against a coronavirus is provided in other aspects of the disclosure. The method involves administering to the subject a composition comprising an mRNA comprising an open reading frame encoding a coronavirus antigen, thereby inducing in the subject an immune response specific to the coronavirus, wherein the immune response in the subject is equivalent to an immune response in a subject vaccinated with a traditional vaccine against the coronavirus at 2 times to 100 times the dosage level relative to the composition.

In some embodiments, the immune response in the subject is equivalent to an immune response in a subject vaccinated with a traditional vaccine at twice the dosage level relative to a composition of the present disclosure. In some embodiments, the immune response in the subject is equivalent to an immune response in a subject vaccinated with a traditional vaccine at three times the dosage level relative to a composition of the present disclosure. In some embodiments, the immune response in the subject is equivalent to an immune response in a subject vaccinated with a traditional vaccine at 4 times, 5 times, 10 times, 50 times, or 100 times the dosage level relative to a composition of the present disclosure. In some embodiments, the immune response in the subject is equivalent to an immune response in a subject vaccinated with a traditional vaccine at 10 times to 1000 times the dosage level relative to a composition of the present disclosure. In some embodiments, the immune response in the subject is equivalent to an immune response in a subject vaccinated with a traditional vaccine at 100 times to 1000 times the dosage level relative to a composition of the present disclosure.

In other embodiments, the immune response is assessed by determining [protein] antibody titer in the subject. In other embodiments, the ability of serum or antibody from an immunized subject is tested for its ability to neutralize viral uptake or reduce coronavirus transformation of human B lymphocytes. In other embodiments, the ability to promote a robust T cell response(s) is measured using art recognized techniques.

Other aspects the disclosure provide methods of eliciting an immune response in a subject against a coronavirus by administering to the subject composition comprising an mRNA having an open reading frame encoding a coronavirus antigen, thereby inducing in the subject an immune response specific to the coronavirus antigen, wherein the immune response in the subject is induced 2 days to 10 weeks earlier relative to an immune response induced in a subject vaccinated with a prophylactically effective dose of a traditional vaccine against the coronavirus. In some embodiments, the immune response in the subject is induced in a subject vaccinated with a prophylactically effective dose of a traditional vaccine at 2 times to 100 times the dosage level relative to a composition of the present disclosure.

In some embodiments, the immune response in the subject is induced 2 days, 3 days, 1 week, 2 weeks, 3 weeks, 5 weeks, or 10 weeks earlier relative to an immune response induced in a subject vaccinated with a prophylactically effective dose of a traditional vaccine.

Also provided herein are methods of eliciting an immune response in a subject against a coronavirus by administering to the subject an mRNA having an open reading frame encoding a first antigen, wherein the RNA does not include a stabilization element, and wherein an adjuvant is not co-formulated or co-administered with the vaccine.

A composition may be administered by any route that results in a therapeutically effective outcome. These include, but are not limited, to intradermal, intramuscular, intranasal, and/or subcutaneous administration. The present disclosure provides methods comprising administering RNA vaccines to a subject in need thereof. The exact amount required will vary from subject to subject, depending on the species, age, and general condition of the subject, the severity of the disease, the particular composition, its mode of administration, its mode of activity, and the like. The RNA is typically formulated in dosage unit form for ease of administration and uniformity of dosage. It will be understood, however, that the total daily usage of the RNA may be decided by the attending physician within the scope of sound medical judgment. The specific therapeutically effective, prophylactically effective, or appropriate imaging dose level for any particular patient will depend upon a variety of factors including the disorder being treated and the severity of the disorder; the activity of the specific compound employed; the specific composition employed; the age, body weight, general health, sex and diet of the patient; the time of administration, route of administration, and rate of excretion of the specific compound employed; the duration of the treatment; drugs used in combination or coincidental with the specific compound employed; and like factors well known in the medical arts.

The effective amount of the RNA, as provided herein, may be as low as 20 μg, administered for example as a single dose or as two 10 μg doses. In some embodiments, the

effective amount is a total dose of 20 μg-300 μg or 25 μg-300 μg. For example, the effective amount may be a total dose of 20 μg, 25 μg, 30 μg, 35 μg, 40 μg, 45 μg, 50 μg, 55 μg, 60 μg, 65 μg, 70 μg, 75 μg, 80 μg, 85 μg, 90 μg, 95 μg, 100 μg, 110 μg, 120 μg, 130 μg, 140 μg, 150 μg, 160 μg, 170 μg, 180 μg, 190 μg, 200 μg, 250 μg, or 300 μg. In some embodiments, the effective amount is a total dose of 20 μg. In some embodiments, the effective amount is a total dose of 25 μg. In some embodiments, the effective amount is a total dose of 50 μg. In some embodiments, the effective amount is a total dose of 75 μg. In some embodiments, the effective amount is a total dose of 100 μg. In some embodiments, the effective amount is a total dose of 150 μg. In some embodiments, the effective amount is a total dose of 200 μg. In some embodiments, the effective amount is a total dose of 250 μg. In some embodiments, the effective amount is a total dose of 300 μg.

The RNA described herein can be formulated into a dosage form described herein, such as an intranasal, intratracheal, or injectable (e.g., intravenous, intraocular, intravitreal, intramuscular, intradermal, intracardiac, intraperitoneal, and subcutaneous).

Vaccine Efficacy

Some aspects of the present disclosure provide formulations of the compositions (e.g., RNA vaccines), wherein the RNA is formulated in an effective amount to produce an antigen specific immune response in a subject (e.g., production of antibodies specific to a coronavirus antigen). "An effective amount" is a dose of the RNA effective to produce an antigen-specific immune response. Also provided herein are methods of inducing an antigen-specific immune response in a subject.

As used herein, an immune response to a vaccine or LNP of the present disclosure is the development in a subject of a humoral and/or a cellular immune response to a (one or more) coronavirus protein(s) present in the vaccine. For purposes of the present disclosure, a "humoral" immune response refers to an immune response mediated by antibody molecules, including, e.g., secretory (IgA) or IgG molecules, while a "cellular" immune response is one mediated by T-lymphocytes (e.g., CD4+ helper and/or CD8+ T cells (e.g., CTLs) and/or other white blood cells. One important aspect of cellular immunity involves an antigen-specific response by cytolytic T-cells (CTLs). CTLs have specificity for peptide antigens that are presented in association with proteins encoded by the major histocompatibility complex (MHC) and expressed on the surfaces of cells. CTLs help induce and promote the destruction of intracellular microbes or the lysis of cells infected with such microbes. Another aspect of cellular immunity involves and antigen-specific response by helper T-cells. Helper T-cells act to help stimulate the function and focus

the activity nonspecific effector cells against cells displaying peptide antigens in association with MHC molecules on their surface. A cellular immune response also leads to the production of cytokines, chemokines, and other such molecules produced by activated T-cells and/or other white blood cells including those derived from CD4+ and CD8+ T-cells.

In some embodiments, the antigen-specific immune response is characterized by measuring an anti-coronavirus antigen antibody titer produced in a subject administered a composition as provided herein. An antibody titer is a measurement of the amount of antibodies within a subject, for example, antibodies that are specific to a particular antigen or epitope of an antigen. Antibody titer is typically expressed as the inverse of the greatest dilution that provides a positive result. Enzyme-linked immunosorbent assay (ELISA) is a common assay for determining antibody titers, for example.

A variety of serological tests can be used to measure antibody against encoded antigen of interest, for example, SAR-CoV-2 virus or SAR-CoV-2 viral antigen, e.g., SAR-CoV-2 spike or S protein, of domain thereof. These tests include the hemagglutination-inhibition test, complement fixation test, fluorescent antibody test, enzyme-linked immunosorbent assay (ELISA), and plaque reduction neutralization test (PRNT). Each of these tests measures different antibody activities. In exemplary embodiments, A plaque reduction neutralization test, or PRNT (e.g., PRNT50 or PRNT90) is used as a serological correlate of protection. PRNT measures the biological parameter of in vitro virus neutralization and is the most serologically virus-specific test among certain classes of viruses, correlating well to serum levels of protection from virus infection.

The basic design of the PRNT allows for virus-antibody interaction to occur in a test tube or microtiter plate, and then measuring antibody effects on viral infectivity by plating the mixture on virus-susceptible cells, preferably cells of mammalian origin. The cells are overlaid with a semi-solid media that restricts spread of progeny virus. Each virus that initiates a productive infection produces a localized area of infection (a plaque), that can be detected in a variety of ways. Plaques are counted and compared back to the starting concentration of virus to determine the percent reduction in total virus infectivity. In PRNT, the serum sample being tested is usually subjected to serial dilutions prior to mixing with a standardized amount of virus. The concentration of virus is held constant such that, when added to susceptible cells and overlaid with semi-solid media, individual plaques can be discerned and counted. In this way, PRNT end-point titers can be calculated for each serum sample at any selected percent reduction of virus activity.

In functional assays intended to assess vaccinal immunogenicity, the serum sample dilution series for antibody titration should ideally start below the "seroprotective" threshold titer. Regarding SARS-CoV-2 neutralizing antibodies, the "seroprotective" threshold titer remains unknown; but a seropositivity threshold of 1:10 can be considered a seroprotection threshold in certain embodiments.

PRNT end-point titers are expressed as the reciprocal of the last serum dilution showing the desired percent reduction in plaque counts. The PRNT titer can be calculated based on a 50% or greater reduction in plaque counts (PRNT50). A PRNT50 titer is preferred over titers using higher cut-offs (e.g., PRNT90) for vaccine sera, providing more accurate results from the linear portion of the titration curve.

There are several ways to calculate PRNT titers. The simplest and most widely used way to calculate titers is to count plaques and report the titer as the reciprocal of the last serum dilution to show >50% reduction of the input plaque count as based on the back-titration of input plaques. Use of curve fitting methods from several serum dilutions may permit calculation of a more precise result. There are a variety of computer analysis programs available for this (e.g., SPSS or GraphPad Prism).

In some embodiments, an antibody titer is used to assess whether a subject has had an infection or to determine whether immunizations are required. In some embodiments, an antibody titer is used to determine the strength of an autoimmune response, to determine whether a booster immunization is needed, to determine whether a previous vaccine was effective, and to identify any recent or prior infections. In accordance with the present disclosure, an antibody titer may be used to determine the strength of an immune response induced in a subject by a composition (e.g., RNA vaccine).

In some embodiments, an anti-coronavirus antigen antibody titer produced in a subject is increased by at least 1 log relative to a control. For example, anti-coronavirus antigen antibody titer produced in a subject may be increased by at least 1.5, at least 2, at least 2.5, or at least 3 log relative to a control. In some embodiments, the anti-coronavirus antigen antibody titer produced in the subject is increased by 1, 1.5, 2, 2.5 or 3 log relative to a control. In some embodiments, the anti-coronavirus antigen antibody titer produced in the subject is increased by 1-3 log relative to a control. For example, the anti-coronavirus antigen antibody titer produced in a subject may be increased by 1-1.5, 1-2, 1-2.5, 1-3, 1.5-2, 1.5-2.5, 1.5-3, 2-2.5, 2-3, or 2.5-3 log relative to a control.

In some embodiments, the anti-coronavirus antigen antibody titer produced in a subject is increased at least 2 times relative to a control. For example, the anti-coronavirus antigen antibody

titer produced in a subject may be increased at least 3 times, at least 4 times, at least 5 times, at least 6 times, at least 7 times, at least 8 times, at least 9 times, or at least 10 times relative to a control. In some embodiments, the anti-coronavirus antigen antibody titer produced in the subject is increased 2, 3, 4, 5, 6, 7, 8, 9, or 10 times relative to a control. In some embodiments, the anti-coronavirus antigen antibody titer produced in a subject is increased 2-10 times relative to a control. For example, the anti-coronavirus antigen antibody titer produced in a subject may be increased 2-10, 2-9, 2-8, 2-7, 2-6, 2-5, 2-4, 2-3, 3-10, 3-9, 3-8, 3-7, 3-6, 3-5, 3-4, 4-10, 4-9, 4-8, 4-7, 4-6, 4-5, 5-10, 5-9, 5-8, 5-7, 5-6, 6-10, 6-9, 6-8, 6-7, 7-10, 7-9, 7-8, 8-10, 8-9, or 9-10 times relative to a control.

In some embodiments, an antigen-specific immune response is measured as a ratio of geometric mean titer (GMT), referred to as a geometric mean ratio (GMR), of serum neutralizing antibody titers to coronavirus. A geometric mean titer (GMT) is the average antibody titer for a group of subjects calculated by multiplying all values and taking the nth root of the number, where n is the number of subjects with available data.

A control, in some embodiments, is an anti-coronavirus antigen antibody titer produced in a subject who has not been administered a composition (e.g., RNA vaccine). In some embodiments, a control is an anti-coronavirus antigen antibody titer produced in a subject administered a recombinant or purified protein vaccine. Recombinant protein vaccines typically include protein antigens that either have been produced in a heterologous expression system (e.g., bacteria or yeast) or purified from large amounts of the pathogenic organism.

In some embodiments, the ability of a composition (e.g., RNA vaccine) to be effective is measured in a murine model. For example, a composition may be administered to a murine model and the murine model assayed for induction of neutralizing antibody titers. Viral challenge studies may also be used to assess the efficacy of a vaccine of the present disclosure. For example, a composition may be administered to a murine model, the murine model challenged with virus, and the murine model assayed for survival and/or immune response (e.g., neutralizing antibody response, T cell response (e.g., cytokine response)).

In some embodiments, an effective amount of a composition (e.g., RNA vaccine) is a dose that is reduced compared to the standard of care dose of a recombinant protein vaccine. A "standard of care," as provided herein, refers to a medical or psychological treatment guideline and can be general or specific. "Standard of care" specifies appropriate treatment based on scientific evidence and collaboration between medical professionals involved in the treatment of a given condition. It is the diagnostic and treatment process that a physician/ clinician should follow for a certain type of patient, illness or clinical circumstance. A "standard of care dose," as

provided herein, refers to the dose of a recombinant or purified protein vaccine, or a live attenuated or inactivated vaccine, or a VLP vaccine, that a physician/clinician or other medical professional would administer to a subject to treat or prevent coronavirus infection or a related condition, while following the standard of care guideline for treating or preventing coronavirus infection or a related condition.

In some embodiments, the anti-coronavirus antigen antibody titer produced in a subject administered an effective amount of a composition is equivalent to an anti-coronavirus antigen antibody titer produced in a control subject administered a standard of care dose of a recombinant or purified protein vaccine, or a live attenuated or inactivated vaccine, or a VLP vaccine.

Vaccine efficacy may be assessed using standard analyses (see, e.g., Weinberg et al., J Infect Dis. 2010 Jun 1;201(11):1607-10). For example, vaccine efficacy may be measured by double-blind, randomized, clinical controlled trials. Vaccine efficacy may be expressed as a proportionate reduction in disease attack rate (AR) between the unvaccinated (ARU) and vaccinated (ARV) study cohorts and can be calculated from the relative risk (RR) of disease among the vaccinated group with use of the following formulas:

Efficacy = (ARU – ARV)/ARU x 100; and

Efficacy = (1-RR) x 100.

Likewise, vaccine effectiveness may be assessed using standard analyses (see, e.g., Weinberg et al., J Infect Dis. 2010 Jun 1;201(11):1607-10). Vaccine effectiveness is an assessment of how a vaccine (which may have already proven to have high vaccine efficacy) reduces disease in a population. This measure can assess the net balance of benefits and adverse effects of a vaccination program, not just the vaccine itself, under natural field conditions rather than in a controlled clinical trial. Vaccine effectiveness is proportional to vaccine efficacy (potency) but is also affected by how well target groups in the population are immunized, as well as by other non-vaccine-related factors that influence the 'real-world' outcomes of hospitalizations, ambulatory visits, or costs. For example, a retrospective case control analysis may be used, in which the rates of vaccination among a set of infected cases and appropriate controls are compared. Vaccine effectiveness may be expressed as a rate difference, with use of the odds ratio (OR) for developing infection despite vaccination:

Effectiveness = (1 – OR) x 100.

In some embodiments, efficacy of the composition (e.g., RNA vaccine) is at least 60% relative to unvaccinated control subjects. For example, efficacy of the composition may be at

least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 95%, at least 98%, or 100% relative to unvaccinated control subjects.

Sterilizing Immunity. Sterilizing immunity refers to a unique immune status that prevents effective pathogen infection into the host. In some embodiments, the effective amount of a composition of the present disclosure is sufficient to provide sterilizing immunity in the subject for at least 1 year. For example, the effective amount of a composition of the present disclosure is sufficient to provide sterilizing immunity in the subject for at least 2 years, at least 3 years, at least 4 years, or at least 5 years. In some embodiments, the effective amount of a composition of the present disclosure is sufficient to provide sterilizing immunity in the subject at an at least 5-fold lower dose relative to control. For example, the effective amount may be sufficient to provide sterilizing immunity in the subject at an at least 10-fold lower, 15-fold, or 20-fold lower dose relative to a control.

Detectable Antigen. In some embodiments, the effective amount of a composition of the present disclosure is sufficient to produce detectable levels of coronavirus antigen as measured in serum of the subject at 1-72 hours post administration.

Titer. An antibody titer is a measurement of the number of antibodies within a subject, for example, antibodies that are specific to a particular antigen (e.g., an anti-coronavirus antigen). Antibody titer is typically expressed as the inverse of the greatest dilution that provides a positive result. Enzyme-linked immunosorbent assay (ELISA) is a common assay for determining antibody titers, for example.

In some embodiments, the effective amount of a composition of the present disclosure is sufficient to produce a 1,000-10,000 neutralizing antibody titer produced by neutralizing antibody against the coronavirus antigen as measured in serum of the subject at 1-72 hours post administration. In some embodiments, the effective amount is sufficient to produce a 1,000-5,000 neutralizing antibody titer produced by neutralizing antibody against the coronavirus antigen as measured in serum of the subject at 1-72 hours post administration. In some embodiments, the effective amount is sufficient to produce a 5,000-10,000 neutralizing antibody titer produced by neutralizing antibody against the coronavirus antigen as measured in serum of the subject at 1-72 hours post administration.

In some embodiments, the neutralizing antibody titer is at least 100 NT_{50}. For example, the neutralizing antibody titer may be at least 200, 300, 400, 500, 600, 700, 800, 900 or 1000 NT_{50}. In some embodiments, the neutralizing antibody titer is at least 10,000 NT_{50}.

In some embodiments, the neutralizing antibody titer is at least 100 neutralizing units per milliliter (NU/mL). For example, the neutralizing antibody titer may be at least 200, 300, 400,

500, 600, 700, 800, 900 or 1000 NU/mL. In some embodiments, the neutralizing antibody titer is at least 10,000 NU/mL.

In some embodiments, an anti-coronavirus antigen antibody titer produced in the subject is increased by at least 1 log relative to a control. For example, an anti-coronavirus antigen antibody titer produced in the subject may be increased by at least 2, 3, 4, 5, 6, 7, 8, 9 or 10 log relative to a control.

In some embodiments, an anti-coronavirus antigen antibody titer produced in the subject is increased at least 2 times relative to a control. For example, an anti-coronavirus antigen antibody titer produced in the subject is increased by at least 3, 4, 5, 6, 7, 8, 9 or 10 times relative to a control.

In some embodiments, a geometric mean, which is the nth root of the product of n numbers, is generally used to describe proportional growth. Geometric mean, in some embodiments, is used to characterize antibody titer produced in a subject.

A control may be, for example, an unvaccinated subject, or a subject administered a live attenuated viral vaccine, an inactivated viral vaccine, or a protein subunit vaccine.

Additional Embodiments

Additional embodiments of the present disclosure are encompassed by the following numbered paragraphs:

1. A messenger ribonucleic acid (mRNA) comprising an open reading frame encoding a fusion protein comprising a receptor binding domain (RBD) of a SARS-CoV-2 Spike protein and a protein transmembrane domain.

2. The mRNA of paragraph 1, wherein the protein transmembrane domain is an influenza hemagglutinin transmembrane domain.

3. The mRNA of paragraph 2, wherein the fusion protein comprises an amino acid sequence having at least 80% identity to the amino acid sequence of SEQ ID NO: 77.

4. The mRNA of paragraph 3, wherein the fusion protein comprises an amino acid sequence having at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 77.

5. The mRNA of paragraph 4, wherein the fusion protein comprises the amino acid sequence of SEQ ID NO: 77.

6. The mRNA of any one of the preceding paragraphs, wherein the open reading frame comprises a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 76.

7. The mRNA of paragraph 6, wherein the open reading frame comprises a nucleotide sequence having at least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 76.

8. The mRNA of paragraph 7, wherein the open reading frame comprises the nucleotide sequence of SEQ ID NO: 76.

9. A messenger ribonucleic acid (mRNA) comprising an open reading frame encoding a fusion protein comprising an amino (N)-terminal domain (NTD) of a SARS-CoV-2 Spike protein and a transmembrane domain.

10. The mRNA of paragraph 9, wherein the transmembrane domain is an influenza hemagglutinin transmembrane domain.

11. The mRNA of paragraph 10, wherein the fusion protein comprises an amino acid sequence having at least 80% identity to the amino acid sequence of SEQ ID NO: 47.

12. The mRNA of paragraph 11, wherein the fusion protein comprises an amino acid sequence having at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 47.

13. The mRNA of paragraph 12, wherein the fusion protein comprises the amino acid sequence of SEQ ID NO: 47.

14. The mRNA of any one of the preceding paragraphs, wherein the open reading frame comprises a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 46.

15. The mRNA of paragraph 14, wherein the open reading frame comprises a nucleotide sequence having at least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 46.

16. The mRNA of paragraph 15, wherein the open reading frame comprises the nucleotide sequence of SEQ ID NO: 46.

17. A messenger ribonucleic acid (mRNA) comprising an open reading frame encoding a fusion protein comprising an amino (N)-terminal domain of a SARS-CoV-2 Spike protein linked to a receptor binding domain of a SARS-CoV-2 Spike protein.

18. The mRNA of paragraph 17, wherein the fusion protein further comprises a transmembrane domain.

19. The mRNA of paragraph 18, wherein the fusion protein comprises an amino acid sequence having at least 80% identity to the amino acid sequence of SEQ ID NO: 92.

20. The mRNA of paragraph 18, wherein the fusion protein comprises an amino acid sequence having at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 92.

21. The mRNA of paragraph 20, wherein the fusion protein comprises the amino acid sequence of SEQ ID NO: 92.

22. The mRNA of any one of the preceding paragraphs, wherein the open reading frame comprises a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 91.

23. The mRNA of paragraph 22, wherein the open reading frame comprises a nucleotide sequence having at least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 91.

24. The mRNA of paragraph 23, wherein the open reading frame comprises the nucleotide sequence of SEQ ID NO: 91.

25. The mRNA of any one of the preceding paragraphs further comprising a 5′ untranslated region (UTR), optionally comprising the nucleotide sequence of SEQ ID NO: 131 or 2.

26. The mRNA of any one of the preceding paragraphs further comprising a 3′ untranslated region (UTR), optionally comprising the nucleotide sequence of SEQ ID NO: 132 or 4.

27. The mRNA of any one of the preceding paragraphs further comprising a 5′ cap, optionally 7mG(5')ppp(5')NlmpNp.

28. The mRNA of any one of the preceding paragraphs further comprising a polyA tail, optionally having a length of about 100 nucleotides.

29. The mRNA of any one of the preceding paragraphs, wherein the mRNA comprises a chemical modification, optionally 1-methylpseudouridine.

30. A composition comprising the mRNA of any one of paragraphs 1-29.

31. A composition comprising the mRNA of any one of paragraphs 1-8 and the mRNA of any one of paragraphs 9-16.

32. A composition comprising the mRNA of any one of paragraphs 17-29.

33. A composition comprising:

(a) a messenger ribonucleic acid (mRNA) comprising an open reading frame encoding a fusion protein comprising a receptor binding domain (RBD) of a SARS-CoV-2 Spike protein and a protein transmembrane domain; and

(b) an mRNA comprising an open reading frame encoding a fusion protein comprising an amino (N)-terminal domain of a SARS-CoV-2 Spike protein and a transmembrane domain.

34. The composition of paragraph 33 wherein the protein transmembrane domain is an influenza hemagglutinin transmembrane domain.

35. The composition of paragraph 34, wherein the fusion protein of (a) comprises an amino acid sequence having at least 80% identity to the amino acid sequence of SEQ ID NO: 77.

36. The composition of paragraph 35, wherein the fusion protein of (a) comprises an amino acid sequence having at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 77.

37. The composition of paragraph 36, wherein the fusion protein of (a) comprises the amino acid sequence of SEQ ID NO: 77.

38. The composition of any one of paragraphs 34-37, wherein the open reading frame of (a) comprises a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 76.

39. The composition of paragraph 38 wherein the open reading frame of (a) comprises a nucleotide sequence having at least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 76.

40. The composition of paragraph 39, wherein the open reading frame of (a) comprises the nucleotide sequence of SEQ ID NO: 76.

41. The composition of any one of paragraphs 34-40, wherein the fusion protein of (b) comprises an amino acid sequence having at least 80% identity to the amino acid sequence of SEQ ID NO: 47.

42. The composition of paragraph 41, wherein the fusion protein of (b) comprises an amino acid sequence having at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 47.

43. The composition of paragraph 42, wherein the fusion protein of (b) comprises the amino acid sequence of SEQ ID NO: 47.

44. The composition of any one of paragraphs 34-43, wherein the open reading frame of (b) comprises a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 46.

45. The composition of paragraph 44, wherein the open reading frame of (b) comprises a nucleotide sequence having at least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 46.

46. The composition of paragraph 45, wherein the open reading frame of (b) comprises the nucleotide sequence of SEQ ID NO: 46.

47. The composition of any one of paragraphs 33-46, wherein the ratio of the mRNA of (a) to the mRNA of (b) is about 1:1.

48. The mRNA of any one of paragraphs 1-29 formulated in a lipid nanoparticle.

49. The composition of any one of paragraphs 30-47 further comprising a lipid nanoparticle.

50. The composition of paragraph 49, wherein the mRNA is formulated in the lipid nanoparticle.

51. The composition of any one of paragraphs 33-47, wherein the mRNA of (a) is formulated in a lipid nanoparticle and the mRNA of (b) is formulated in a lipid nanoparticle.

52. The composition of paragraph 51, wherein the mRNA of (a) and (b) are in the same lipid nanoparticle or wherein each of the mRNA of (a) and (b) is formulated in a separate nanoparticle, relative to each other.

53. The mRNA of paragraph 48 or the composition of any one of paragraphs 49-52, wherein the lipid nanoparticle comprises a cationic lipid.

54. The mRNA or composition of paragraph 53, wherein the lipid nanoparticle further comprises a neutral lipid.

55. The mRNA or composition of paragraph 53 or 54, wherein the lipid nanoparticle further comprises a sterol.

56. The mRNA or composition of any one of paragraphs 53-55, wherein the lipid nanoparticle further comprises a polyethylene glycol (PEG)-modified lipid.

57. The mRNA or composition of any one of paragraphs 53-56, wherein the lipid nanoparticle comprises an ionizable cationic lipid, a neutral lipid, a sterol, and a PEG-modified lipid.

58. The mRNA or composition of paragraph 57, wherein the ionizable cationic lipid is heptadecan-9-yl 8 ((2 hydroxyethyl)(6 oxo 6-(undecyloxy)hexyl)amino)octanoate (Compound 1).

59. The mRNA or composition of paragraph 57 or 58, wherein the neutral lipid is 1,2 distearoyl-sn-glycero-3-phosphocholine (DSPC).

60. The mRNA or composition of any one of paragraphs 57-59, wherein the sterol is cholesterol.

61. The mRNA or composition of any one of paragraphs 57-60, wherein the PEG-modified lipid is 1,2 dimyristoyl-sn-glycerol, methoxypolyethyleneglycol (PEG2000 DMG).

62. The mRNA or composition of any one of paragraphs 57-61, wherein the lipid nanoparticle comprises 20-60 mol% ionizable cationic lipid, 5-25 mol% neutral lipid, 25-55 mol% sterol, and 0.5-15 mol% PEG-modified lipid.

63. The mRNA or composition of paragraph 62, wherein the lipid nanoparticle comprises:

47 mol% ionizable cationic lipid; 11.5 mol% neutral lipid; 38.5 mol% sterol; and 3.0 mol% PEG-modified lipid;

48 mol% ionizable cationic lipid; 11 mol% neutral lipid; 38.5 mol% sterol; and 2.5 mol% PEG-modified lipid;

49 mol% ionizable cationic lipid; 10.5 mol% neutral lipid; 38.5 mol% sterol; and 2.0 mol% PEG-modified lipid;

50 mol% ionizable cationic lipid; 10 mol% neutral lipid; 38.5 mol% sterol; and 1.5 mol% PEG-modified lipid; or

51 mol% ionizable cationic lipid; 9.5 mol% neutral lipid; 38.5 mol% sterol; and 1.0 mol% PEG-modified lipid.

64. The mRNA or composition of paragraph 63, wherein the lipid nanoparticle comprises:

47 mol% Compound 1; 11.5 mol% DSPC; 38.5 mol% cholesterol; and 3.0 mol% PEG2000 DMG;

48 mol% Compound 1; 11 mol% DSPC; 38.5 mol% cholesterol; and 2.5 mol% PEG2000 DMG;

49 mol% Compound 1; 10.5 mol% DSPC; 38.5 mol% cholesterol; and 2.0 mol% PEG2000 DMG;

50 mol% Compound 1; 10 mol% DSPC; 38.5 mol% cholesterol; and 1.5 mol% PEG2000 DMG; or

51 mol% Compound 1; 9.5 mol% DSPC; 38.5 mol% cholesterol; and 1.0 mol% PEG2000 DMG.

65. A method comprising administering to a subject the mRNA or the composition of any one of the preceding paragraphs in an amount effective to induce in the subject a neutralizing antibody response against SARS-CoV-2.

66. A method comprising administering to a subject the mRNA or the composition of any one of the preceding paragraphs in an amount effective to induce in the subject a T cell immune response against SARS-CoV-2.

67. A messenger ribonucleic acid (mRNA) comprising an open reading frame (ORF) that encodes a coronavirus antigen capable of inducing an immune response, such as a neutralizing antibody response, to a SARS-CoV-2, wherein the antigen comprises a protein fragment or a functional protein domain of a SARS-CoV-2, optionally wherein the RNA is formulated in a lipid nanoparticle.

68. The mRNA of paragraph 67, wherein the antigen is a functional protein domain.

69. The mRNA of paragraph 68, wherein the protein domain is an N-terminal domain (NTD) of a SARS-CoV-2 Spike protein.

70. The mRNA of paragraph 69, wherein the NTD is linked to a transmembrane domain, optionally an influenza hemagglutinin transmembrane domain.

71. The mRNA of paragraph 70, wherein the antigen comprises an amino acid sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 47, optionally wherein the antigen comprises the amino acid sequence of SEQ ID NO: 47.

72. The mRNA of paragraph 70 or 71, wherein the open reading frame comprises a nucleotide sequence having at least 70%, least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 46, optionally wherein the open reading frame comprises the nucleotide sequence of SEQ ID NO: 46.

73. The mRNA of paragraph 68, wherein the protein domain is a receptor binding domain (RBD) of a SARS-CoV-2 Spike protein.

74. The mRNA of paragraph 73, wherein the RBD is soluble.

75. The mRNA of paragraph 74, wherein the antigen comprises an amino acid sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 62, optionally wherein the antigen comprises the amino acid sequence of SEQ ID NO: 62.

76. The mRNA of paragraph 74 or 75, wherein the open reading frame comprises a nucleotide sequence having at least 70%, least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 61, optionally wherein the open reading frame comprises the nucleotide sequence of SEQ ID NO: 61.

77. The mRNA of paragraph 73, wherein the RBD is linked to a transmembrane domain, optionally an influenza hemagglutinin transmembrane domain.

78. The mRNA of paragraph 77, wherein the antigen comprises an amino acid sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 77, optionally wherein the antigen comprises the amino acid sequence of SEQ ID NO: 77.

79. The mRNA of paragraph 77 or 78, wherein the open reading frame comprises a nucleotide sequence having at least 70%, least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ

ID NO: 76, optionally wherein the open reading frame comprises the nucleotide sequence of SEQ ID NOs: 76.

80. The mRNA of paragraph 69, wherein the NTD is linked to an RBD of a SARS-CoV-2 Spike protein to form an NTD-RBD fusion protein.

81. The mRNA of paragraph 80, wherein the NTD-RBD fusion is linked to a transmembrane domain (TM), optionally an influenza hemagglutinin transmembrane domain, to form an NTD-RBD-TM protein.

82. The mRNA of paragraph 81, wherein the antigen comprises an amino acid sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 92, optionally wherein the antigen comprises the amino acid sequence of SEQ ID NO: 92.

83. The mRNA of paragraph 81 or 82, wherein the open reading frame comprises a nucleotide sequence having at least 70%, least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 91, optionally wherein the open reading frame comprises the nucleotide sequence of SEQ ID NO: 91.

84. The mRNA of paragraph 80, wherein the NTD-RBD fusion comprises a C-terminal truncation.

85. The mRNA of paragraph 84, wherein the antigen comprises an amino acid sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 107, optionally wherein the antigen comprises the amino acid sequence of SEQ ID NO: 107.

86. The mRNA of paragraph 84 or 85, wherein the open reading frame comprises a nucleotide sequence having at least 70%, least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 106, optionally wherein the open reading frame comprises the nucleotide sequence of SEQ ID NO: 106.

87. The mRNA of any one of the preceding paragraphs, wherein the NTD and/or RBD includes an extended region.

88. The mRNA of paragraph 87, wherein the antigen comprises an amino acid sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of any one of SEQ ID NOs: 59, 86, 89, 116, 119, or 122, optionally wherein the antigen comprises the amino acid sequence of any one of SEQ ID NOs: 59, 86, 89, 116, 119, or 122.

89. The mRNA of paragraph 87 or 88, wherein the open reading frame comprises a nucleotide sequence having at least 70%, least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of any one of SEQ ID NOs: 58, 85, 88, 115, 118, or 121, optionally wherein the open reading frame comprises the nucleotide sequence of any one of SEQ ID NOs: 58, 85, 88, 115, 118, or 121.

90. The mRNA of paragraph 68, wherein the protein domain is an S1 subunit domain of a SARS-CoV-2 Spike protein.

91. The mRNA of paragraph 90, wherein the S1 subunit is soluble.

92. The mRNA of paragraph 91, wherein the antigen comprises an amino acid sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 5, optionally wherein the antigen comprises the amino acid sequence of SEQ ID NO: 5.

93. The mRNA of paragraph 91 or 92, wherein the open reading frame comprises a nucleotide sequence having at least 70%, least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 3, optionally wherein the open reading frame comprises the nucleotide sequence of SEQ ID NO: 3.

94. The mRNA of paragraph 90, wherein the S1 subunit is linked to a transmembrane domain, optionally an influenza hemagglutinin transmembrane domain.

95. The mRNA of paragraph 94, wherein the antigen comprises an amino acid sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 17, optionally wherein the antigen comprises the amino acid sequence of SEQ ID NO: 17.

96. The mRNA of paragraph 94 or 95, wherein the open reading frame comprises a nucleotide sequence having at least 70%, least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 16, optionally wherein the open reading frame comprises the nucleotide sequence of SEQ ID NO: 16.

97. The mRNA of paragraph 90, wherein the S1 subunit has been modified to remove a RBD or a portion of a RBD of S protein.

98. The mRNA of paragraph 97, wherein the antigen comprises an amino acid sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of any one of SEQ ID NOs: 20, 23, 26,

29, 32 or 35, optionally wherein the antigen comprises the amino acid sequence of any one of SEQ ID NOs: 20, 23, 26, 29, 32, or 35.

99. The mRNA of paragraph 97 or 98, wherein the open reading frame comprises a nucleotide sequence having at least 70%, least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of any one of SEQ ID NOs: 19, 22, 25, 28, 41, or 34, optionally wherein the open reading frame comprises the nucleotide sequence of any one of SEQ ID NOs: 19, 22, 25, 28, 31, or 34.

100. The mRNA of paragraph 90, wherein the S1 subunit is linked to an S2 subunit of an S protein.

101. The mRNA of paragraph 100, wherein the S2 subunit is from a SARS-CoV-2 S protein and in some embodiments wherein the S2 subunit comprises an open reading frame comprising a nucleotide sequence having at least 70%, least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 145, optionally wherein the open reading frame comprises the nucleotide sequence of SEQ ID NOs: 145.

102. The mRNA of paragraph 101, wherein the S1 subunit is from an HKU1 S protein.

103. The mRNA of paragraph 102, wherein the antigen comprises an amino acid sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 38, optionally wherein the antigen comprises the amino acid sequence of SEQ ID NO: 38.

104. The mRNA of paragraph 102 or 103, wherein the open reading frame comprises a nucleotide sequence having at least 70%, least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 37, optionally wherein the open reading frame comprises the nucleotide sequence of SEQ ID NO: 37.

105. The mRNA of paragraph 101, wherein the S1 subunit is from an OC43 S protein.

106. The mRNA of paragraph 105, wherein the antigen comprises an amino acid sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 41, optionally wherein the antigen comprises the amino acid sequence of SEQ ID NO: 41.

107. The mRNA of paragraph 105 or 106, wherein the open reading frame comprises a nucleotide sequence having at least 70%, least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ

ID NO: 40, optionally wherein the open reading frame comprises the nucleotide sequence of SEQ ID NO: 40.

108. The mRNA of any one of the preceding paragraphs, wherein the antigen further comprises a scaffold domain, optionally selected from ferritin, lumazine synthetase and a foldon.

109. The mRNA of paragraph 108, wherein the scaffold domain is ferritin.

110. The mRNA of paragraph 109, wherein the antigen comprises an amino acid sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 8 or 65, optionally wherein the antigen comprises the amino acid sequence of SEQ ID NO: 8 or 65.

111. The mRNA of paragraph 109 or 110, wherein the open reading frame comprises a nucleotide sequence having at least 70%, least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 7 or 64, optionally wherein the open reading frame comprises the nucleotide sequence of SEQ ID NO: 7 or 64.

112. The mRNA of paragraph 108, wherein the scaffold domain is lumazine synthetase.

113. The mRNA of paragraph 112, wherein the antigen comprises an amino acid sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of any one of SEQ ID NOs: 11, 14, 68, or 71, optionally wherein the antigen comprises the amino acid sequence of any one of SEQ ID NOs: 11, 14, 68, or 71.

114. The mRNA of paragraph 112 or 113, wherein the open reading frame comprises a nucleotide sequence having at least 70%, least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of any one of SEQ ID NOs: 10, 13, 67, or 70, optionally wherein the open reading frame comprises the nucleotide sequence of any one of SEQ ID NOs: 10, 13, 67, or 70.

115. The mRNA of paragraph 108, wherein the scaffold domain is a foldon.

116. The mRNA of paragraph 115, wherein the antigen comprises an amino acid sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of any one of SEQ ID NOs: 44, 50, 74, 80, 83, 101, 104 or 113, optionally wherein the antigen comprises the amino acid sequence of any one of SEQ ID NOs: 44, 50, 74, 80, 83, 101, 104 or 113.

117. The mRNA of paragraph 115 or 116, wherein the open reading frame comprises a nucleotide sequence having at least 70%, least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of any

one of SEQ ID NOs: 43, 49, 73, 79, 82, 100, 103, or 112, optionally wherein the open reading frame comprises the nucleotide sequence of any one of SEQ ID NOs: 43, 49, 73, 79, 82, 100, 103, or 112.

118. The mRNA of any one of the preceding paragraphs, wherein the antigen further comprising a trafficking signal, optionally selected from macrophage markers, optionally CD86, CD11B and/or VSVGct.

119. The mRNA of paragraph 118, wherein the antigen comprises an amino acid sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of any one of SEQ ID NOs: 95, 98, or 110, optionally wherein the antigen comprises the amino acid sequence of any one of SEQ ID NOs: 95, 98, or 110.

120. The mRNA of paragraph 118 or 119, wherein the open reading frame comprises a nucleotide sequence having at least 70%, least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of any one of SEQ ID NOs: 94, 97, or 109, optionally wherein the open reading frame comprises the nucleotide sequence of any one of SEQ ID NOs: 94, 97, or 109.

121. The mRNA of any one of paragraphs 67-120 formulated in a lipid nanoparticle.

122. The mRNA of paragraph 121, wherein the lipid nanoparticle comprises a cationic lipid, optionally an ionizable cationic lipid, a neutral lipid, a sterol, and/or a polyethylene glycol (PEG)-modified lipid.

123. The mRNA or composition of paragraph 108, wherein the ionizable cationic lipid is heptadecan-9-yl 8 ((2 hydroxyethyl)(6 oxo 6-(undecyloxy)hexyl)amino)octanoate (Compound 1), the neutral lipid is 1,2 distearoyl-sn-glycero-3-phosphocholine (DSPC), the sterol is cholesterol, and/or the PEG-modified lipid is 1,2 dimyristoyl-sn-glycerol, methoxypolyethyleneglycol (PEG2000 DMG).

124. The mRNA or composition of any one of paragraphs 121-123, wherein the lipid nanoparticle comprises 20-60 mol% ionizable cationic lipid, 5-25 mol% neutral lipid, 25-55 mol% sterol, and 0.5-15 mol% PEG-modified lipid.

125. The mRNA or composition of paragraph 124, wherein the lipid nanoparticle comprises:

47 mol% ionizable cationic lipid; 11.5 mol% neutral lipid; 38.5 mol% sterol; and 3.0 mol% PEG-modified lipid;

48 mol% ionizable cationic lipid; 11 mol% neutral lipid; 38.5 mol% sterol; and 2.5 mol% PEG-modified lipid;

49 mol% ionizable cationic lipid; 10.5 mol% neutral lipid; 38.5 mol% sterol; and 2.0 mol% PEG-modified lipid;

50 mol% ionizable cationic lipid; 10 mol% neutral lipid; 38.5 mol% sterol; and 1.5 mol% PEG-modified lipid; or

51 mol% ionizable cationic lipid; 9.5 mol% neutral lipid; 38.5 mol% sterol; and 1.0 mol% PEG-modified lipid.

126. The mRNA or composition of paragraph 125, wherein the lipid nanoparticle comprises:

47 mol% Compound 1; 11.5 mol% DSPC; 38.5 mol% cholesterol; and 3.0 mol% PEG2000 DMG;

48 mol% Compound 1; 11 mol% DSPC; 38.5 mol% cholesterol; and 2.5 mol% PEG2000 DMG;

49 mol% Compound 1; 10.5 mol% DSPC; 38.5 mol% cholesterol; and 2.0 mol% PEG2000 DMG;

50 mol% Compound 1; 10 mol% DSPC; 38.5 mol% cholesterol; and 1.5 mol% PEG2000 DMG; or

51 mol% Compound 1; 9.5 mol% DSPC; 38.5 mol% cholesterol; and 1.0 mol% PEG2000 DMG.

127. A method comprising administering to a subject the mRNA of any one of the paragraphs 67-126 in an amount effective to induce in the subject a neutralizing antibody response against SARS-CoV-2.

128. A method comprising administering to a subject the mRNA of any one of the paragraphs 67-126 in an amount effective to induce in the subject a T cell immune response against SARS-CoV-2.

EXAMPLES

Example 1. Expression Data

The mRNAs used in the present study were used to express key neutralizing domains of the SARS-CoV-2 coronavirus spike (S) protein and assess whether these neutralizing protein domains may be more efficient at inducing protective immunity when used individually or in combination as an immunogenic composition or vaccine to protect people from infection by the live and spreading natural virus. The linear designs of the proteins encoded by the mRNAs are shown in **FIG. 2**. The proteins all also contain a carboxy (C) terminal transmembrane domain (TM) derived from the hemagglutinin (HA) of influenza.

Both the NTD and RBD are known to be sites for binding of antibodies that manifest neutralizing virus activity. RBD in the case of SARS-CoV-2 is the receptor binding site of the spike protein and binds the angiotensin-converting enzyme 2 (ACE2). The amino (N) terminal domain, NTD, the function of which is not thoroughly understood, seems to have a role in binding sugar moieties and in facilitating the conformational transition of the spike protein from prefusion to a post fusion conformation. *See* Zhou H, Chen Y, Zhang S, *et al. Nat Commun.* 2019; 10(1): 3068. Regardless, both the NTD and RBD domains induce high binding antibody and neutralizing antibody titers as discussed below.

Expression data for the mRNA RBD-TM vaccine ("SARS-CoV-2 RBD-TM"; SEQ ID NOs: 75-77), the mRNA NTD-TM vaccine ("SARS-CoV-2 NTD-TM"; SEQ ID NOs: 45-47), and mRNA NTD-RBD-TM ("SARS-CoV-1 NTD-RBD-TM"; SEQ ID NOs: 90-92) vaccine is shown in **Tables 16 and 17** using an antibody specific for the receptor binding domain (RBD) of SARS-CoV-2 Spike protein (mAb1) and for the N-terminal domain (NTD) of SARS-CoV-2 Spike protein (Ab2). **Table 16** shows the average fold difference (over dilution range) in MFI*Freq compared to WT SARS-CoV-2 Spike protein mRNA (**Table 16**) at 24 hours (hr), 48 hr, and 72 hr.

Table 16. Total Antigen Expression (MFI * Freq) Fold Change Compared to S2P Protein

	Monoclonal Antibody (mAb)	**Wild Type (WT) S mRNA**	**RBD-TM mRNA**	**NTD-TM mRNA**	**NTD-RBD-TM mRNA**	**blank**
24 hour (hr)	1 – specific to RBD	1.8	30.6	0.0	10.4	0.0
24 hr	2 – specific to NTD	0.7	0.0	9.7	2.5	0.0
48 hr	1	1.2	40.1	0.0	7.3	0.0
48 hr	2	0.8	0.0	19.0	6.0	0.0
72 hr	1	0.8	32.5	0.1	11.6	0.0
72 hr	2	0.7	0	14.4	4.1	0.0

RBD=receptor binding domain
NTD=N-terminal domain
TM=transmembrane domain

Example 2. Immunogenicity Data and Neutralization Data at Day 21 Following a Single Dose

mRNA NTD-TM and mRNA RBD-TM (described in **Example 1**) were administered to mice at the following doses: 0.001 μg, 0.01 μg, 0.1 μg, or 1 μg (N=8). mRNA NTD-RBD-TM (described in **Example 1**) were administered to mice at the following doses: 0.1 μg or 1 μg (N=8). A 50:50 mixture of the mRNA NTD-TM and mRNA RBD-TM was administered to mice, which contained 0.1 μg of each mRNA for a total of 0.2 μg mRNA, or 1 μg of each mRNA

for a total of 2 μg mRNA (N=8). SARS-CoV-2 Spike protein-specific IgG titers (**Table 17**), SARS-CoV-2 RBD-specific IgG titers (**Table 18**), and SARS-CoV-2 NTD-specific IgG titers (**Table 19**) were then measured by ELISA at Day 21 post vaccination. The data is provided in **Tables 17-19**. The 0.1 μg dose of mRNA NTD-RBD-TM and 0.2 μg dose of a 50:50 mixture of mRNA NTD-TM and mRNA RBD-TM compositions, elicited observable NTD-specific and RBD-specific IgG titers, and the 0.1 μg doses of RBD-TM and NTD-TM elicited measurable IgG titers against RBD and NTD antigens, respectively.

Table 17. SARS-CoV-2 S1/S2 spike protein-specific IgG titers – Mean Values

Micrograms (per mRNA construct)	RBD-TM	NTD-TM	NTD-RBD-TM	50:50 mix
0.001	13	13	-	-
0.01	13	13	-	-
0.1	13	13	13	19
1	225	13	379	707

Average of N=8; PBS only = 1.1

Table 18. RBD domain-specific IgG titers – Mean Values

Micrograms (per mRNA construct)	RBD-TM	NTD-RBD-TM	50:50 mix
0.001	15	-	-
0.01	14	-	-
0.1	75	209	266
1	1947	7710	15530

Average of N=8; PBS only = 1.1

Table 19. NTD domain-specific IgG titers – Mean Values

Micrograms (per mRNA construct)	NTD-TM	NTD-RBD-TM	50:50 mix
0.001	13	-	-
0.01	14	-	-
0.1	30	20	38
1	620	848	1075

Average of N=8; PBS only = 1.1

Neutralization titers from serum of mice vaccinated with the 1 μg dose of RBD-TM and NTD-RBD-TM compositions and the 2 μg dose of the 50:50 mixture of NTD-TM and RBD-TM

compositions were measured and the correlation between ELISA titers and neutralization titers was analyzed (**FIG. 7**).

The titers elicited by 1 μg dose of the NTD-RBD-TM composition or 2 μg of the 50:50 mixture of NTD-TM and RBD-TM compositions were greater than those elicited by the 1 μg dose of RBD-TM composition (**Table 20**). Significant correlations exist between neutralization titers and ELISA titers of Spike-specific IgG, RBD-specific IgG, and NTD-specific IgG (**FIG. 7**).

Table 20. Neutralization Titers – Mean Values

Micrograms (per mRNA construct)	RBD-TM	NTD-RBD-TM	50:50 mix
1	238	504	-
2	-	-	421

Average of N=8; PBS only = 1.1

***Recombinant VSVΔG*-based SARS-CoV-2 Pseudovirus Neutralization Assay**

Codon-optimized wild-type or D614G spike gene (Wuhan-Hu-1 strain; NC_045512.2) was cloned into pCAGGS vector. To generate VSVΔG-based SARS-CoV-2 pseudovirus, BHK-21/WI-2 cells were transfected with the spike expression plasmid and infected VSVΔG-firefly-luciferase as previously described (Whitt, 2010). A549-hACE2-TMPRSS2 cells were used as target cells for VSVΔG-based SARS-CoV-2 pseudovirus neutralization assay. Lentivirus encoding hACE2-P2A-TMPRSS2 was made to generate A549-hACE2-TMPRSS2 cells which were maintained in DMEM supplemented with 10% fetal bovine serum and 1 μg/mL puromycin. A549-hACE2-TMPRSS2 cells were infected by pseudovirus for 1 hr at 37 Celsius. The inoculum virus or virus-antibody mix was removed after infection. 18 hr later, equal volume of One-Glo reagent (Promega; E6120) was added to culture medium for readout using BMG PHERastar-FS plate reader. The neutralization procedure and data analysis are same as mentioned above in the lentivirus-based pseudovirus neutralization assay. See Whitt, M.A. (2010). Journal of Virological Methods 169, 365–374.

Example 3. Immunogenicity Data at Day 36 Following Two Doses

The same doses of the mRNA vaccines described in **Example 2** were again administered to mice as booster doses on Day 22 post-vaccination with the first dose. The titers of antibodies generated after the booster dose to each of RBD antigen, NTD antigen, wildtype (WT) Spike (S) protein and S2P protein (S protein having a double proline mutations to stabilize the prefusion

conformation) were measured by ELISA from day 36 serum and shown below. The 50:50 mixture of the two immunogenic compositions of RBD-TM and NTD-TM encoded by mRNA in an LNP were administered at 2 μg or 0.2 μg total mRNA to mice as a booster dose on day 22 and the titers were determined on day 36. See **Table 21**.

The WT S protein titers shown in **Table 21** by mice immunized with RBD-TM, NTD-TM, or the NTD-RBD-TM encoded by mRNA in an LNP indicated that two doses were superior at all doses tested in inducing antibodies that could recognize and bind to SARS-CoV-2 WT S protein.

Table 21. SARS-CoV-2 WT S-specific IgG titers – Geometric Mean Values

Micrograms (mRNA construct)	RBD-TM	NTD-TM	NTD-RBD-TM
	Day 21 (GMT) Titers		
0.001	13	13	-
0.01	13	13	-
0.1	13	13	14
1	225	13	379
	Day 36 (GMT) Titers		
0.001	13	13	-
0.01	295	13	-
0.1	1,017	39	3879
1	27,674	6317	24,413

Average of N=8; PBS only = 1.1

The serum from mice immunized with two doses of RBD-TM, NTD-TM, or the NTD-RBD-TM encoded by mRNA in an LNP and were further analyzed for the ability of the antibodies to recognize and bind SARS-CoV-2 S2P protein. The titers to SARS-CoV-2 S2P protein were determined by ELISA using the S2P as the antigen on the plate and are shown below in **Table 22**. Each of these immunogens induced much higher antibody titers when S2P was the antigen versus when WT S protein was the ELISA antigen. Compare **Table 21** and **Table 22**.

Table 22. SARS-CoV-2 S2P-specific IgG titers – Geometric Mean Values

Micrograms (mRNA construct)	RBD-TM	NTD-TM	NTD-RBD-TM
	Day 36 (GMT) Titers		
0.001	88	16	-
0.01	1,416	280	-
0.1	3,892	11,687	25,298
1	107,611	192,040	223,140

Average of N=8; PBS only = 1.1

In **Table 23**, the immunogen was the 50:50 mix of RBD-TM and NTD-TM encoded by mRNA in an LNP and the titers to WT S, RBD, NTD, and S2P were determined after one dose (day 21) and two doses (day 36). These results show dramatically increased titers when the immunogen is the combination of RBD-TM and NTD-TM in a 50:50 mix compared to the antibody titers induced by the individual antigens at the same doses. The 50:50 mix induced good titers to the immunizing antigens, but surprisingly even better titers to the WT S protein and extremely higher titers to the S2P protein. See **Table 23**.

Table 23. SARS-CoV-2 Antigen-specific IgG titers – Geometric Mean Values

Immunogen	Detection antigen on ELISA plate			
Micrograms (mRNA) 50:50 mix RBD-TM LNP NTD-TM LNP	Anti-Spike Protein IgG	Anti-RBD IgG	Anti-NTD IgG	Anti-S2P IgG
Day 21 (GMT) Titers				
0.2	19	266	38	4,869
2	707	15,530	1,075	52,936
Day 36 (GMT) Titers				
0.2	7,259	15,957	4,869	74,718
2	75,871	186,740	52,936	2,726,948

Average of N=8; PBS only = 1.1

Table 24 shows the results of an immunization with each of RBD-TM and NTD-TM as mRNA encoding those antigens. The geometric mean titers were measured for groups of 8 mice using the protein encoded by the mRNA immunogen as the antigen on the ELISA plate. Again, both immunogenic compositions induced high titers to the immunizing antigen when the antigen is administered as mRNA formulated in an LNP. In this case, two doses produced superior antibody responses at all concentrations. However, on a per microgram (μg) basis the 50:50 mix of these antigens induced about a 10-fold higher antibody response than when the antigens are administered singly. Compare **Table 23** and **Table 24**.

Table 24. SARS-CoV-2 RBD and NTD Domain Specific IgG titers – Geometric Mean Values

Micrograms (mRNA construct)	RBD-TM Antigen on plate	NTD-TM Antigen on plate
	Day 21 (GMT) Titers	
0.001	15	13
0.01	14	14
0.1	75	30
1	1,947	620
	Day 36 (GMT) Titers	

0.001	25	13
0.01	1,357	123
0.1	4,678	6898
1	84,385	46,516

Average of N=8; PBS only = 1.1

The fusion protein comprising NTD linked to RBD and encoded by mRNA in an LNP was administered as an immunogenic composition to groups of 8 mice at 0.1 and 1 μg doses at days 1 and 21. See **Table 25** below. Even the mRNA encoding the fusion protein version of NTD-RBD-TM induced very good titers to the individual domains that were higher than when a single domain was the immunizing antigen. The titers to the S2P protein were about 8-fold higher than the titers to WT S protein. See **Table 25**.

Table 25. SARS-CoV-2 NTD-RBD-TM Domain Specific IgG titers – Geometric Mean Values

Micrograms Immunogen NTD-RBD-TM (mRNA construct)	RBD-TM Antigen on plate	NTD-TM Antigen on plate	WT Spike Antigen on plate	S2P Antigen on plate
		Day 21 (GMT) Titers		
0.1	209	20	13	NA
1	7,710	848	379	NA
		Day 36 (GMT) Titers		
0.1	5,769	2,875	3,878	25,298
1	111,747	41,876	24,413	223,140

Average of N=8; PBS only = 1.1

Neutralization data are shown in **Table 26**. The S1-666-TM encoded by mRNA is an antigen using the S1 subdomain, specifically residues 1-666 of SARS-CoV-2 spike protein attached to a transmembrane domain.

Table 26. Mean Neutralization Titers on Day 36 after Two Immunizations on Days 1 and 22.

Micrograms (per mRNA construct)	S1-666-TM	50:50 Mix of NTD-TM & RBD-TM	RBD-TM	NTD-TM	NTD-RBD-TM
0.001	-	-	40	40	-
0.01	508	-	70.5	40	-
0.1	11,561	1,336	260	205	865
1.0	-	25,208	7,402	4,669	57,891

Average of N=8; PBS only = 1.1

Example 4. Immunogenicity of S1-666-TM

The S1-666-TM (or S1 residues 1-666 of spike protein S) encoded by mRNA in an LNP was administered to mice as a prime immunization on day 1, and as booster dose on day 22 for 0.01 μg , and 0.1 μg (N=8) groups. The titers of antibodies generated after the booster dose to each of mRNA RBD, mRNA NTD, and mRNA wildtype (WT) Spike (S) protein (**FIG. 1**) were measured by ELISA from days 21 (pre-boost) and 36 (post boost) serum and shown below in **Table 27**.

The WT S protein titers shown in **Table 27** by mice immunized with S1-666-TM encoded by mRNA in an LNP indicated that two doses were superior at all doses tested in inducing antibodies that could recognize and bind to SARS-CoV-2 WT S protein. Surprisingly, the induced titers weres highest when measured against the S2P version of the spike protein even though the 2P mutation is not found in S1, because the 2P mutation occurs in S2 and S2 is not present in the immunogen. Similar to other constructs, the NTD titers require the second dose to become elevated.

Table 27. SARS-CoV-2 Antigen-specific IgG titers – Geometric Mean Values

Immunogen	Detection antigen on ELISA plate			
Micrograms (mRNA) S1-666-TM	Anti-Spike Protein IgG	Anti-RBD IgG	Anti-NTD IgG	Anti-S2P IgG
Day 21 (GMT) Titers				
0.01	13	30	13	NA
0.1	175	868	49	NA
Day 36 (GMT) Titers				
0.01	2,450	12,952	112	186,561
0.1	17,066	69,969	9,015	874,609

Average of N=8; PBS only = 1.1

Example 5. Immunogenicity of RBD-TM, NTD-TM, NTD-RBD-TM, and 50:50 mixture of NTD-TM/RBD-TM compositions at Day 36 after two doses

In this repeat experiment, the same doses of the mRNA vaccines described in the examples above were again administered to mice as booster doses on Day 22 post-vaccination with the first dose. SARS-CoV-2 Spike protein-specific IgG titers, SARS-CoV-2 S2P protein-specific IgG titers, SARS-CoV-2 RBD-specific IgG titers, and SARS-CoV-2 NTD-specific IgG titers (were then measured by ELISA at Day 36 post-vaccination with the first dose.

Results showed that 1 μg and 0.1 μg doses of mRNA RBD-TM, mRNA NTD-TM, mRNA NTD-RBD-TM compositions, and 50:50 mixtures containing 1 μg or 0.1 μg each of

mRNA RBD-TM and mRNA NTD-TM compositions, elicited high ELISA titers towards SARS-CoV-2 Spike or SARS-CoV-2 S2P proteins.

Example 6. Immunogenicity Studies

The immunogenicity of a 50:50 mixture of mRNA NTD-TM and mRNA RBD-TM was administered to mice at the follow doses: 0.2 μg or 2 μg total mRNA (0.1 μg or 1 μg of each mRNA) (N=8). A prime dose was administered on Day 1, and a boost dose was administered on Day 22. On Day 36, ELISA was used to assess antibody binding to SARS-CoV-2 stabilized prefusion spike protein (SARS-CoV-2 pre-S). The following vaccine compositions including mRNA NTD-RBD-TM, mRNA RBD-TM and mRNA NTD-TM were administered to mice at the following doses: 0.1 μg and 0.01 μg (N=8). The GMT data was determined and shown below in **Table 28**.

Table 28. SARS-CoV-2 Spike Titers induced by S Protein Domain Fusions and Combinations

Vaccination Immunogen	**Dose μg mRNA**	**Day 21 GMT Titer 3 weeks post dose 1 S2P coated on plates**	**Day 36 GMT Titer 3 weeks post dose 2 S2P coated on plates**
50:50 mix RBD-TM NTD-TM	1	5,913	301,881
50:50 mix RBD-TM NTD-TM	0.1	71	13,423
NTD-RBD-TM	1.0	3,329	177,499
NTD-RBD-TM	0.1	80	48947
RBD-TM	1.0	3198	593,730
RBD-TM	0.1	73	31,090
NTD-TM	1.0	593	74,777
NTD-TM	0.1	21	23,916

Example 7. Determination of the Ratio of IgG2a and IgG1 for NTD-RBD-TM

The compositions of NTD-RBD-TM, mRNA NTD-RBD-TM were administered to mice at the following doses: 0.1 μg and 1 μg. A prime dose was administered on Day 1, and a boost dose was administered on Day 22. On Day 36, S2P specific IgG1 and IgG2a titers were assessed. See **FIGs. 8A-8C.** By day 36, the titer of IgG2a was higher than the amount of IgG1 at both dose levels. See **FIG. 8A**. To determine whether the T cell response is skewed toward either a Th1 or a Th2 type of response, we plotted the ration of IgG2a/IgG1 at the day 36 timepoint. As shown in **FIG. 8B**, the NTD-RBD-TM composition induces an antibody

within the Th1 type of response. Th2 type response is disfavored in vaccine development because of an association with driving disease enhancement.

Example 8. Immunogenocity Studies

The mRNAs listed in **Table 29** were administered to mice at the following doses: 0.1 μg and 1 μg (N=8). A prime dose was administered on Day 1, and a boost dose was administered on Day 22. Serum IgG tiers were assayed on S2P coated plates on Day 21 and Day 36. Results are shown in **Table 29**.

Table 29.

Formulation/Material	Dose (μg)	Day 21	Day 36
PBS		1.097	1.097
NTD-RBD-TM	1	4.237	6.036
	0.1	2.711	5.106
S1-666-TM (SEQ ID NO: 16)	1	4.224	5.704
	0.1	1.944	4.633
mix (50:50) NTD-EXT-F43C-TM (SEQ ID NO: 55) + RBD-Q563D-EXT-TM (SEQ ID NO: 88)	1	3.939	5.442
	0.1	1.867	4.342
NTD-EXT-RBD-EXT-TM (SEQ ID NO: 121)	1	4.541	6.001
	0.1	3.694	4.885
S1-594-TM (SEQ ID NO: 22)	1	4.027	5.517
	0.1	3.313	4.780
S1-594-PolyG-DS-TM (SEQ ID NO: 28)	1	3.974	5.531
	0.1	2.551	4.425
S2-TM (SEQ ID NO: 145)	1	1.989	4.058
	0.1	1.097	2.015
RBD-NTD-TM (SEQ ID NO: 139)	1	4.523	6.062
	0.1	3.453	5.222
NTD-PADRE-RBD-TM (SEQ ID NO: 142)	1	4.489	6.040
	0.1	3.386	5.301

ADDITIONAL SEQUENCES

It should be understood that any of the mRNA sequences described herein may include a 5′ UTR and/or a 3′ UTR. The UTR sequences may be selected from the following sequences, or other known UTR sequences may be used. It should also be understood that any of the mRNA constructs described herein may further comprise a poly(A) tail and/or cap (e.g., 7mG(5')ppp(5')NlmpNp). Further, while many of the mRNAs and encoded antigen sequences described herein include a signal peptide and/or a peptide tag (e.g., C-terminal His tag), it should

be understood that the indicated signal peptide and/or peptide tag may be substituted for a different signal peptide and/or peptide tag, or the signal peptide and/or peptide tag may be omitted.

5' UTR: GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGAGCCACC (SEQ ID NO: 131)

5' UTR: GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCGGCGCCGCCACC (SEQ ID NO: 2)

3' UTR:
UGAUAAUAGGCUGGAGCCUCGGUGGCCAUGCUUCUUGCCCCUUGGGCCUCCCCCCAGCCCCUCCUCCCCUUCCUGCAC
CCGUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC (SEQ ID NO: 132)

3' UTR:
UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGGCCUCCCCCCAGCCCCUCCUCCCCUUCCUGCAC
CCGUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC (SEQ ID NO: 4)

SARS-CoV-2 Wild Type Spike (S) Protein		
SEQ ID NO: 123 consists of from 5' end to 3' end: 5' UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 124, and 3' UTR SEQ ID NO: 4.		123
Chemistry	1-methylpseudouridine	
Cap	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCG GCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGUGAGCAGCCAGUGCG UGAACCUGACCACCCGGACCCAGCUGCCACCAGCCUACACCAACAG CUUCACCCGGGGCGUCUACUACCCCGACAAGGUGUUCCGGAGCAGC GUCCUGCACAGCACCCAGGACCUGUUCCUGCCCUUCUUCAGCAACG UGACCUGGUUCCACGCCAUCCACGUGAGCGGCACCAACGGCACCAA GCGGUUCGACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUC GCCAGCACCGAGAAGAGCAACAUCAUCCGGGGCUGGAUCUUCGGCA CCACCCUGGACAGCAAGACCCAGAGCCUGCUGAUCGUGAAUAACGC CACCAACGUGGUGAUCAAGGUGUGCGAGUUCCAGUUCUGCAACGAC CCCUUCCUGGGCGUGUACUACCACAAGAACAACAAGAGCUGGAUGG AGAGCGAGUUCCGGGUGUACAGCAGCGCCAACAACUGCACCUUCGA GUACGUGAGCCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAGGGC AACUUCAAGAACCUGCGGGAGUUCGUGUUCAAGAACAUCGACGGCU ACUUCAAGAUCUACAGCAAGCACACCCCAAUCAACCUGGUGCGGGA UCUGCCCCAGGGCUUCUCAGCCCUGGAGCCCCUGGUGGACCUGCCC AUCGGCAUCAACAUCACCCGGUUCCAGACCCUGCUGGCCCUGCACC GGAGCUACCUGACCCCAGGCGACAGCAGCAGCGGGUGGACAGCAGG CGCGGCUGCUUACUACGUGGGCUACCUGCAGCCCCGGACCUUCCUG CUGAAGUACAACGAGAACGGCACCAUCACCGACGCCGUGGACUGCG CCCUGGACCCUCUGAGCGAGACCAAGUGCACCCUGAAGAGCUUCAC CGUGGAGAAGGGCAUCUACCAGACCAGCAACUUCCGGGUGCAGCCC ACCGAGAGCAUCGUGCGGUUCCCCAACAUCACCAACCUGUGCCCCU UCGGCGAGGUGUUCAACGCCACCCGGUUCGCCAGCGUGUACGCCUG GAACCGGAAGCGGAUCAGCAACUGCGUGGCCGACUACAGCGUGCUG UACAACAGCGCCAGCUUCAGCACCUUCAAGUGCUACGGCGUGAGCC CCACCAAGCUGAACGACCUGUGCUUCACCAACGUGUACGCCGACAG CUUCGUGAUCCGUGGCGACGAGGUGCGGCAGAUCGCACCCGGCCAG	124

	ACAGGCAAGAUCGCCGACUACAACUACAAGCUGCCCGACGACUUCA CCGGCUGCGUGAUCGCCUGGAACAGCAACAACCUCGACAGCAAGGU GGGCGGCAACUACAACUACCUGUACCGGCUGUUCCGGAAGAGCAAC CUGAAGCCCUUCGAGCGGGACAUCAGCACCGAGAUCUACCAAGCCG GCUCCACCCCUUGCAACGGCGUGGAGGGCUUCAACUGCUACUUCCC UCUGCAGAGCUACGGCUUCCAGCCCACCAACGGCGUGGGCUACCAG CCCUACCGGGUGGUGGUGCUGAGCUUCGAGCUGCUGCACGCCCCAG CCACCGUGUGUGGCCCCAAGAAGAGCACCAACCUGGUGAAGAACAA GUGCGUGAACUUCAACUUCAACGGCCUUACCGGCACCGGCGUGCUG ACCGAGAGCAACAAGAAAUUCCUGCCCUUUCAGCAGUUCGGCCGGG ACAUCGCCGACACCACCGACGCUGUGCGGGAUCCCCAGACCCUGGA GAUCCUGGACAUCACCCCUUGCAGCUUCGGCGGCGUGAGCGUGAUC ACCCCAGGCACCAACACCAGCAACCAGGUGGCCGUGCUGUACCAGG ACGUGAACUGCACCGAGGUGCCCGUGGCCAUCCACGCCGACCAGCU GACACCCACCUGGCGGGUCUACAGCACCGGCAGCAACGUGUUCCAG ACCCGGGCCGGUUGCCUGAUCGGCGCCGAGCACGUGAACAACAGCU ACGAGUGCGACAUCCCCAUCGGCGCCGGCAUCUGUGCCAGCUACCA GACCCAGACCAAUUCACCCCGGAGGGCAAGGAGCGUGGCCAGCCAG AGCAUCAUCGCCUACACCAUGAGCCUGGGCGCCGAGAACAGCGUGG CCUACAGCAACAACAGCAUCGCCAUCCCCACCAACUUCACCAUCAG CGUGACCACCGAGAUUCUGCCCGUGAGCAUGACCAAGACCAGCGUG GACUGCACCAUGUACAUCUGCGGCGACAGCACCGAGUGCAGCAACC UGCUGCUGCAGUACGGCAGCUUCUGCACCCAGCUGAACCGGGCCCU GACCGGCAUCGCCGUGGAGCAGGACAAGAACACCCAGGAGGUGUUC GCCCAGGUGAAGCAGAUCUACAAGACCCCUCCCAUCAAGGACUUCG GCGGCUUCAACUUCAGCCAGAUCCUGCCCGACCCCAGCAAGCCCAG CAAGCGGAGCUUCAUCGAGGACCUGCUGUUCAACAAGGUGACCCUA GCCGACGCCGGCUUCAUCAAGCAGUACGGCGACUGCCUCGGCGACA UAGCCGCCCGGGACCUGAUCUGCGCCCAGAAGUUCAACGGCCUGAC CGUGCUGCCUCCCCUGCUGACCGACGAGAUGAUCGCCCAGUACACC AGCGCCCUGUUAGCCGGAACCAUCACCAGCGGCUGGACUUUCGGCG CUGGAGCCGCUCUGCAGAUCCCCUUCGCCAUGCAGAUGGCCUACCG GUUCAACGGCAUCGGCGUGACCCAGAACGUGCUGUACGAGAACCAG AAGCUGAUCGCCAACCAGUUCAACAGCGCCAUCGGCAAGAUCCAGG ACAGCCUGAGCAGCACCGCUAGCGCCCUGGGCAAGCUGCAGGACGU GGUGAACCAGAACGCCCAGGCCCUGAACACCCUGGUGAAGCAGCUG AGCAGCAACUUCGGCGCCAUCAGCAGCGUGCUGAACGACAUCCUGA GCCGGCUGGACAAGGUGGAGGCCGAGGUGCAGAUCGACCGGCUGAU CACUGGCCGGCUGCAGAGCCUGCAGACCUACGUGACCCAGCAGCUG AUCCGGGCCGCCGAGAUUCGGGCCAGCGCCAACCUGGCCGCCACCA AGAUGAGCGAGUGCGUGCUGGGCCAGAGCAAGCGGGUGGACUUCUG CGGCAAGGGCUACCACCUGAUGAGCUUUCCCCAGAGCGCACCCCAC GGAGUGGUGUUCCUGCACGUGACCUACGUGCCCGCCCAGGAGAAGA ACUUCACCACCGCCCCAGCCAUCUGCCACGACGGCAAGGCCCACUU UCCCCGGGAGGGCGUGUUCGUGAGCAACGGCACCCACUGGUUCGUG ACCCAGCGGAACUUCUACGAGCCCCAGAUCAUCACCACCGACAACA CCUUCGUGAGCGGCAACUGCGACGUGGUGAUCGGCAUCGUGAACAA CACCGUGUACGAUCCCCUGCAGCCCGAGCUGGACAGCUUCAAGGAG GAGCUGGACAAGUACUUCAAGAAUCACACCAGCCCCGACGUGGACC UGGGCGACAUCAGCGGCAUCAACGCCAGCGUGGUGAACAUCCAGAA GGAGAUCGAUCGGCUGAACGAGGUGGCCAAGAACCUGAACGAGAGC CUGAUCGACCUGCAGGAGCUGGGCAAGUACGAGCAGUACAUCAAGU GGCCCUGGUACAUCUGGCUGGGCUUCAUCGCCGGCCUGAUCGCCAU CGUGAUGGUGACCAUCAUGCUGUGCUGCAUGACCAGCUGCUGCAGC UGCCUGAAGGGCUGUUGCAGCUGCGGCAGCUGCUGCAAGUUCGACG AGGACGACAGCGAGCCCGUGCUGAAGGGCGUGAAGCUGCACUACAC C	

3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGG CCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGC	4
Corresponding amino acid sequence	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSS VLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYF ASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCND PFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQG NFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLP IGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFL LKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQP TESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVL YNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQ TGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSN LKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQ PYRVVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVL TESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVI TPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQ TRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPRRARSVASQ SIIAYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSV DCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVF AQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTL ADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYT SALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQ KLIANQFNSAIGKIQDSLSSTASALGKLQDVVNQNAQALNTLVKQL SSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTYVTQQL IRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPH GVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFVSNGTHWFV TQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKE ELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVAKNLNES LIDLQELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCS CLKGCCSCGSCCKFDEDDSEPVLKGVKLHYT	125
PolyA tail	100 nt	

It should also be understood that any one of the open reading frames and/or corresponding amino acid sequences described herein may include or exclude a signal sequence.

All references, patents and patent applications disclosed herein are incorporated by reference with respect to the subject matter for which each is cited, which in some cases may encompass the entirety of the document.

The indefinite articles "a" and "an," as used herein in the specification and in the claims, unless clearly indicated to the contrary, should be understood to mean "at least one."

It should also be understood that, unless clearly indicated to the contrary, in any methods claimed herein that include more than one step or act, the order of the steps or acts of the method is not necessarily limited to the order in which the steps or acts of the method are recited.

In the claims, as well as in the specification above, all transitional phrases such as "comprising," "including," "carrying," "having," "containing," "involving," "holding,"

"composed of," and the like are to be understood to be open-ended, i.e., to mean including but not limited to. Only the transitional phrases "consisting of" and "consisting essentially of" shall be closed or semi-closed transitional phrases, respectively, as set forth in the United States Patent Office Manual of Patent Examining Procedures, Section 2111.03.

The terms "about" and "substantially" preceding a numerical value mean ±10% of the recited numerical value.

Where a range of values is provided, each value between the upper and lower ends of the range are specifically contemplated and described herein.

The entire contents of International Application Nos. PCT/US2015/02740, PCT/US2016/043348, PCT/US2016/043332, PCT/US2016/058327, PCT/US2016/058324, PCT/US2016/058314, PCT/US2016/058310, PCT/US2016/058321, PCT/US2016/058297, PCT/US2016/058319, and PCT/US2016/058314 are incorporated herein by reference.

CLAIMS

What is claimed is:

1. A messenger ribonucleic acid (mRNA) comprising an open reading frame (ORF) that encodes at least two domains of a SARS-CoV-2 Spike protein, and less than the full length spike protein.

2. The mRNA of claim 1, wherein one of the two domains is an N-terminal domain (NTD) of a SARS-CoV-2 Spike protein.

3. The mRNA of claim 1 or 2, wherein one of the two domains is a receptor binding domain (RBD) of a SARS-CoV-2 Spike protein.

4. The mRNA of claim 2 or 3, wherein the ORF encodes a transmembrane domain (TD) linked to the NTD and/or RBD.

5. The mRNA of claim 4, wherein the TD is an influenza hemagglutinin transmembrane domain.

6. The mRNA of claim 3 or 4, wherein the ORF comprises NTD - RBD - TM.

7. The mRNA of any one of claims 1-6, wherein the at least two domains are linked through a cleavable or non-cleavable linker.

8. The mRNA of claim 7, wherein the non-cleavable linker is a glycine-serine (GS) linker.

9. The mRNA of claim 8, wherein the GS linker 4-15 amino acids.

10. The mRNA of claim 7, wherein the linker is a pan HLA DR-binding epitope (PADRE).

11. The mRNA of any one of claims 1-10, wherein the ORF encodes a signal peptide.

12. The mRNA of claim 11, wherein the signal peptide is linked to the NTD.

13. The mRNA of claim 11, wherein the signal peptide is linked to the RBD.

14. The mRNA of any one of claims 11-13, wherein the signal peptide is heterologous to SARS-CoV-2.

15. The mRNA of any one of claims 1-3, wherein the at least two domains are soluble.

16. The mRNA of any one of claims 1-14, wherein the ORF encodes a trafficking signal domain.

17. The mRNA of claim 16, wherein the trafficking signal domain is a macrophage marker.

18. The mRNA of claim 17, wherein the macrophage marker CD86 and/or CD11b.

19. The mRNA of claim 16, wherein the trafficking signal domain is a VSV-G cytosolic tail (VSVGct).

20. The mRNA of claim 1, wherein one of the two domains is a first repetitive heptapeptide: HPPHCPC (HR1) of a SARS-CoV-2 Spike protein.

21. The mRNA of claim 1, wherein one of the two domains is a second repetitive heptapeptide: HPPHCPC (HR2) of a SARS-CoV-2 Spike protein.

22. The mRNA of claim 20 or 21, wherein the ORF encodes a transmembrane domain (TD) linked to the HR1 and/or HR2.

23. The mRNA of claim 22, wherein the TD is an influenza hemagglutinin transmembrane domain.

24. The mRNA of any one of claims 20-23, wherein the ORF encodes a fusion peptide (FP).

25. The mRNA of any one of claims 20-23, wherein the ORF encodes a CT tail.

26. A messenger ribonucleic acid (mRNA) comprising an open reading frame (ORF) that encodes a receptor binding domain (RBD) of a SARS-CoV-2 Spike protein.

27. The mRNA of claim 26, wherein the RBD is soluble.

28. The mRNA of claim 26, wherein the RBD is linked to a transmembrane domain, optionally an influenza hemagglutinin transmembrane domain.

29. A messenger ribonucleic acid (mRNA) comprising an open reading frame (ORF) that encodes an N-terminal domain (NTD) of a SARS-CoV-2 Spike protein.

30. The mRNA of claim 29, wherein the NTD is linked to an RBD of a SARS-CoV-2 Spike protein to form an NTD-RBD fusion protein.

31. The mRNA of claim 30, wherein the NTD-RBD fusion is linked to a transmembrane domain (TM), optionally an influenza hemagglutinin transmembrane domain, to form an NTD-RBD-TM protein.

32. The mRNA of claim 30, wherein the NTD-RBD fusion comprises a C-terminal truncation.

33. The mRNA of any one of claims 26-32, wherein the NTD and/or RBD includes an extended region.

34. A messenger ribonucleic acid (mRNA) comprising an open reading frame (ORF) that encodes an S1 subunit, wherein the S1 subunit is linked to a transmembrane domain, optionally an influenza hemagglutinin transmembrane domain.

35. A messenger ribonucleic acid (mRNA) comprising an open reading frame (ORF) that encodes an S1 subunit, wherein the S1 subunit has been modified to remove a RBD or a portion of a RBD of S protein.

36. A messenger ribonucleic acid (mRNA) comprising an open reading frame (ORF) that encodes an S1 subunit linked to an S2 subunit, wherein the S1 subunit is from an HKU1 S protein.

37. A messenger ribonucleic acid (mRNA) comprising an open reading frame (ORF) that encodes an S1 subunit linked to an S2 subunit, wherein the S1 subunit is from an OC43 S protein.

38. The mRNA of any one of claims 34-37, further comprising a trafficking signal, optionally selected from macrophage markers, optionally CD86, CD11B and/or VSVGct.

39. A messenger ribonucleic acid (mRNA) comprising an open reading frame (ORF) that encodes an S2 subunit of a SARS-CoV-2 Spike protein.

40. A messenger ribonucleic acid (mRNA) comprising an open reading frame encoding a fusion protein comprising a receptor binding domain (RBD) of a SARS-CoV-2 Spike protein and a protein transmembrane domain.

41. The mRNA of claim 40, wherein the protein transmembrane domain is an influenza hemagglutinin transmembrane domain.

42. A messenger ribonucleic acid (mRNA) comprising an open reading frame encoding a fusion protein comprising an amino (N)-terminal domain (NTD) of a SARS-CoV-2 Spike protein and a transmembrane domain.

43. The mRNA of claim 42, wherein the transmembrane domain is an influenza hemagglutinin transmembrane domain.

44. A messenger ribonucleic acid (mRNA) comprising an open reading frame encoding a fusion protein comprising an amino (N)-terminal domain of a SARS-CoV-2 Spike protein linked to a receptor binding domain of a SARS-CoV-2 Spike protein.

45. The mRNA of claim 44, wherein the fusion protein further comprises a transmembrane domain.

46. The mRNA of any one of the preceding claims further comprising a 5′ cap, optionally 7mG(5')ppp(5')NlmpNp.

47. The mRNA of any one of claims 1-46, further comprising a polyA tail, optionally having a length of about 100 nucleotides.

48. The mRNA of any one of claims 1-47, wherein the mRNA comprises a chemical modification, optionally 1-methylpseudouridine.

49. The mRNA of any one of claims 1-48, wherein each uridine in the mRNA is a 1-methylpseudouridine.

50. A composition comprising:

(a) a messenger ribonucleic acid (mRNA) comprising an open reading frame encoding a fusion protein comprising a receptor binding domain (RBD) of a SARS-CoV-2 Spike protein and a protein transmembrane domain; and

(b) an mRNA comprising an open reading frame encoding a fusion protein comprising an amino (N)-terminal domain (NTD) of a SARS-CoV-2 Spike protein and a transmembrane domain.

51. The composition of claim 50, wherein the protein transmembrane domain is an influenza hemagglutinin transmembrane domain.

52. The composition of claim 51, wherein the fusion protein of (a) comprises an amino acid sequence having at least 80% identity to the amino acid sequence of SEQ ID NO: 77.

53. The composition of claim 52, wherein the fusion protein of (a) comprises an amino acid sequence having at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 77.

54. The composition of claim 53, wherein the fusion protein of (a) comprises the amino acid sequence of SEQ ID NO: 77.

55. The composition of any one of claims 50-54, wherein the open reading frame of (a) comprises a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 76.

56. The composition of claim 55, wherein the open reading frame of (a) comprises a nucleotide sequence having at least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 76.

57. The composition of claim 56, wherein the open reading frame of (a) comprises the nucleotide sequence of SEQ ID NO: 76.

58. The composition of any one of claims 50-57, wherein the fusion protein of (b) comprises an amino acid sequence having at least 80% identity to the amino acid sequence of SEQ ID NO: 47.

59. The composition of claim 58, wherein the fusion protein of (b) comprises an amino acid sequence having at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 47.

60. The composition of claim 59, wherein the fusion protein of (b) comprises the amino acid sequence of SEQ ID NO: 47.

61. The composition of any one of claims 50-60, wherein the open reading frame of (b) comprises a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 46.

62. The composition of claim 61, wherein the open reading frame of (b) comprises a nucleotide sequence having at least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 46.

63. The composition of claim 62, wherein the open reading frame of (b) comprises the nucleotide sequence of SEQ ID NO: 46.

64. The composition of any one of claims 50-63, wherein the ratio of the mRNA of (a) to the mRNA of (b) is about 1:1.

65. The mRNA of any one of claims 1-49, in a composition further comprising a lipid nanoparticle.

66. The composition of any one of claims 50-64 further comprising a lipid nanoparticle.

67. The composition of claim 65 or 66, wherein the mRNA is in the lipid nanoparticle.

68. The composition of any one of claims 50-64, wherein the mRNA of (a) is in a lipid nanoparticle and the mRNA of (b) is in a lipid nanoparticle.

69. The composition of claim 68, wherein the mRNA of (a) and (b) are in the same lipid nanoparticle or wherein each of the mRNA of (a) and (b) is formulated in a separate nanoparticle, relative to each other.

70. The composition of any one of claims 65-69, wherein the lipid nanoparticle comprises a cationic lipid.

71. The composition of claim 70, wherein the lipid nanoparticle further comprises a neutral lipid.

72. The composition of claim 70 or 71, wherein the lipid nanoparticle further comprises a sterol.

73. The composition of any one of claims 70-72, wherein the lipid nanoparticle further comprises a polyethylene glycol (PEG)-modified lipid.

74. The composition of any one of claims 70-73, wherein the lipid nanoparticle comprises an ionizable cationic lipid, a neutral lipid, a sterol, and a PEG-modified lipid.

75. The composition of claim 74, wherein the ionizable cationic lipid is heptadecan-9-yl 8 ((2 hydroxyethyl)(6 oxo 6-(undecyloxy)hexyl)amino)octanoate (Compound 1).

76. The composition of claim 74 or 75, wherein the neutral lipid is 1,2 distearoyl-sn-glycero-3-phosphocholine (DSPC).

77. The composition of any one of claims 70-76, wherein the sterol is cholesterol.

78. The composition of any one of claims 70-77, wherein the PEG-modified lipid is 1,2 dimyristoyl-sn-glycerol, methoxypolyethyleneglycol (PEG2000 DMG).

79. The composition of any one of claims 70-78, wherein the lipid nanoparticle comprises 20-60 mol% ionizable cationic lipid, 5-25 mol% neutral lipid, 25-55 mol% sterol, and 0.5-15 mol% PEG-modified lipid.

80. The composition of claim 79, wherein the lipid nanoparticle comprises:
40-50 mol% ionizable lipid, optionally 45-50 mol%, 30-45 mol% sterol, optionally 35-40 mol%, 5-15 mol% helper lipid, optionally 10-12 mol%, 1-5% PEG lipid, optionally 1-3 mol%, or 1.5 to 2.5 mol%.

81. The composition of claim 80, wherein the lipid nanoparticle comprises:
40-50 mol% Compound 1, optionally 45-50 mol%, 30-45 mol% cholesterol, optionally 35-40 mol%, 5-15 mol% DSPC, optionally 10-12 mol%, 1-5% PEG2000DMG, optionally 1-3 mol%, or 1.5 to 2.5 mol%.

82. A method comprising administering to a subject the mRNA or the composition of any one of the preceding claims in an amount effective to induce in the subject a neutralizing antibody response against SARS-CoV-2.

83. A method comprising administering to a subject the mRNA or the composition of any one of the preceding claims in an amount effective to induce in the subject a T cell immune response against SARS-CoV-2.

84. A method comprising administering to a subject the mRNA of any one of the claims 1-49 in an amount effective to induce in the subject a neutralizing antibody response and a T cell immune response against SARS-CoV-2.

85. A messenger ribonucleic acid (mRNA) comprising an open reading frame (ORF) that encodes at least one domain of a SARS-CoV-2 Spike protein.

86. The mRNA of claim 85, wherein the mRNA is any one of claims 1-49.

87. The mRNA of claim 85 or 86, wherein the ORF encodes an amino acid sequence having at least 80% identity to the amino acid sequence of SEQ ID NO: 77.

88. The mRNA of claim 85 or 86, wherein the ORF encodes an amino acid sequence having at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 77.

89. The mRNA of claim 85 or 86, wherein the ORF encodes an amino acid sequence of SEQ ID NO: 77.

90. The mRNA of claim 85 or 86, wherein the ORF comprises a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 76.

91. The mRNA of claim 85 or 86, wherein the ORF comprises a nucleotide sequence having at least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 76.

92. The mRNA of claim 85 or 86, wherein the ORF comprises the nucleotide sequence of SEQ ID NO: 76.

93. The mRNA of claim 85 or 86, wherein the ORF encodes an amino acid sequence having at least 80% identity to the amino acid sequence of SEQ ID NO: 47.

94. The mRNA of claim 85 or 86, wherein the ORF encodes an amino acid sequence having at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 47.

95. The mRNA of claim 85 or 86, wherein the ORF encodes the amino acid sequence of SEQ ID NO: 47.

96. The mRNA of claim 85 or 86, wherein the ORF comprises a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 46.

97. The mRNA of claim 85 or 86, wherein the ORF comprises a nucleotide sequence having at least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 46.

98. The mRNA of claim 85 or 86, wherein the ORF comprises the nucleotide sequence of SEQ ID NO: 46.

99. The mRNA of claim 85 or 86, wherein the ORF encodes an amino acid sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 62, optionally wherein the antigen comprises the amino acid sequence of SEQ ID NO: 62.

100. The mRNA of claim 85 or 86, wherein the ORF comprises a nucleotide sequence having at least 70%, at least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 61, optionally wherein the open reading frame comprises the nucleotide sequence of SEQ ID NO: 61.

101. The mRNA of claim 85 or 86, wherein the ORF encodes an amino acid sequence having at least 80% identity to the amino acid sequence of SEQ ID NO: 92, 140, or 143.

102. The mRNA of claim 85 or 86, wherein the ORF encodes an amino acid sequence having at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 92, 140, or 143.

103. The mRNA of claim 85 or 86, wherein the ORF encodes the amino acid sequence of SEQ ID NO: 92, 140, or 143.

104. The mRNA of claim 85 or 86, wherein the ORF comprises a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 91, 139, or 142.

105. The mRNA of claim 85 or 86, wherein the ORF comprises a nucleotide sequence having at least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 91, 139, or 142.

106. The mRNA of claim 85 or 86, wherein the ORF comprises the nucleotide sequence of SEQ ID NO: 91, 139, or 142.

107. The mRNA of claim 85 or 86, wherein the ORF encodes an amino acid sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 107, optionally wherein the antigen comprises the amino acid sequence of SEQ ID NO: 107.

108. The mRNA of claim 85 or 86, wherein the ORF comprises a nucleotide sequence having at least 70%, at least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 106, optionally wherein the open reading frame comprises the nucleotide sequence of SEQ ID NO: 106.

109. The mRNA of claim 85 or 86, wherein the ORF encodes an amino acid sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of any one of SEQ ID NOs: 59, 86, 89, 116, 119, or 122, optionally wherein the antigen comprises the amino acid sequence of any one of SEQ ID NOs: 59, 86, 89, 116, 119, or 122.

110. The mRNA of claim 85 or 86, wherein the ORF comprises a nucleotide sequence having at least 70%, at least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of any one of SEQ ID NOs: 58, 85, 88, 115, 118, or 121, optionally wherein the open reading frame comprises the nucleotide sequence of any one of SEQ ID NOs: 58, 85, 88, 115, 118, or 121.

111. A messenger ribonucleic acid (mRNA) comprising an open reading frame (ORF) that encodes an S1 subunit antigen that comprises an amino acid sequence having at least 80%, at

least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 5, optionally wherein the antigen comprises the amino acid sequence of SEQ ID NO: 5.

112. The mRNA of claim 85 or 86, wherein the ORF comprises a nucleotide sequence having at least 70%, at least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 3, optionally wherein the open reading frame comprises the nucleotide sequence of SEQ ID NO: 3.

113. The mRNA of claim 85 or 86, wherein the antigen comprises an amino acid sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 17 or 146, optionally wherein the antigen comprises the amino acid sequence of SEQ ID NO: 17 or 146.

114. The mRNA of claim 85 or 86, wherein the ORF comprises a nucleotide sequence having at least 70%, at least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 16 or 145, optionally wherein the open reading frame comprises the nucleotide sequence of SEQ ID NO: 16 or 145.

115. The mRNA of claim 85 or 86, wherein the ORF encodes an amino acid sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of any one of SEQ ID NOs: 20, 23, 26, 29, 32 or 35, optionally wherein the antigen comprises the amino acid sequence of any one of SEQ ID NOs: 20, 23, 26, 29, 32, or 35.

116. The mRNA of claim 85 or 86, wherein the ORF comprises a nucleotide sequence having at least 70%, at least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of any one of SEQ ID NOs: 19, 22, 25, 28, 31, or 34, optionally wherein the open reading frame comprises the nucleotide sequence of any one of SEQ ID NOs: 19, 22, 25, 28, 31, or 34.

117. The mRNA of claim 85 or 86, wherein the ORF encodes an amino acid sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or

at least 99% identity to the amino acid sequence of SEQ ID NO: 38, optionally wherein the antigen comprises the amino acid sequence of SEQ ID NO: 38.

118. The mRNA of claim 85 or 86, wherein the ORF comprises a nucleotide sequence having at least 70%, at least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 37, optionally wherein the open reading frame comprises the nucleotide sequence of SEQ ID NO: 37.

119. The mRNA of claim 85 or 86, wherein the ORF encodes an amino acid sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 41, optionally wherein the antigen comprises the amino acid sequence of SEQ ID NO: 41.

120. The mRNA of claim 85 or 86, wherein the ORF comprises a nucleotide sequence having at least 70%, at least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 40, optionally wherein the open reading frame comprises the nucleotide sequence of SEQ ID NO: 40.

121. The mRNA of any one of the preceding claims further comprising a 5′ untranslated region (UTR), optionally comprising the nucleotide sequence of SEQ ID NO: 131 or 2.

122. The mRNA of any one of the preceding claims further comprising a 3′ untranslated region (UTR), optionally comprising the nucleotide sequence of SEQ ID NO: 132 or 4.

123. A messenger ribonucleic acid (mRNA) comprising an open reading frame (ORF) that encodes an antigen that comprises an amino acid sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 8 or 65, optionally wherein the antigen comprises the amino acid sequence of SEQ ID NO: 8 or 65.

124. The mRNA of claim 123, wherein the ORF comprises a nucleotide sequence having at least 70%, at least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 7 or 64, optionally wherein the open reading frame comprises the nucleotide sequence of SEQ ID NO: 7 or 64.

125. A messenger ribonucleic acid (mRNA) comprising an open reading frame (ORF) that encodes an antigen that comprises an amino acid sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of any one of SEQ ID NOs: 11, 14, 68, or 71, optionally wherein the antigen comprises the amino acid sequence of any one of SEQ ID NOs: 11, 14, 68, or 71.

126. The mRNA of claim 125, wherein the open reading frame comprises a nucleotide sequence having at least 70%, at least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of any one of SEQ ID NOs: 10, 13, 67, or 70, optionally wherein the open reading frame comprises the nucleotide sequence of any one of SEQ ID NOs: 10, 13, 67, or 70.

127. A messenger ribonucleic acid (mRNA) comprising an open reading frame (ORF) that encodes an antigen that comprises an amino acid sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of any one of SEQ ID NOs: 44, 50, 74, 80, 83, 101, 104 or 113, optionally wherein the antigen comprises the amino acid sequence of any one of SEQ ID NOs: 44, 50, 74, 80, 83, 101, 104 or 113.

128. The mRNA of claim 127, wherein the open reading frame comprises a nucleotide sequence having at least 70%, at least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of any one of SEQ ID NOs: 43, 49, 73, 79, 82, 100, 103, or 112, optionally wherein the open reading frame comprises the nucleotide sequence of any one of SEQ ID NOs: 43, 49, 73, 79, 82, 100, 103, or 112.

129. The mRNA of claim 109, wherein the antigen comprises an amino acid sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of any one of SEQ ID NOs: 95, 98, or 110, optionally wherein the antigen comprises the amino acid sequence of any one of SEQ ID NOs: 95, 98, or 110.

130. The mRNA of claim 129, wherein the open reading frame comprises a nucleotide sequence having at least 70%, at least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of any one of SEQ ID NOs: 94, 97, or 109, optionally wherein the open reading frame comprises the nucleotide sequence of any one of SEQ ID NOs: 94, 97, or 109.

131. The mRNA of any one of claims 85-130, in a composition further comprising a lipid nanoparticle.

132. The composition of claim 131, wherein the mRNA is in the lipid nanoparticle.

133. The composition of any one of claims 131-132, wherein the lipid nanoparticle comprises a cationic lipid.

134. The composition of claim 133, wherein the lipid nanoparticle further comprises a neutral lipid.

135. The composition of any one of claims 131-134, wherein the lipid nanoparticle further comprises a sterol.

136. The composition of any one of claims 131-135, wherein the lipid nanoparticle further comprises a polyethylene glycol (PEG)-modified lipid.

137. The composition of any one of claims 131-136, wherein the lipid nanoparticle comprises an ionizable cationic lipid, a neutral lipid, a sterol, and a PEG-modified lipid.

138. The composition of claim 137, wherein the ionizable cationic lipid is heptadecan-9-yl 8 ((2 hydroxyethyl)(6 oxo 6-(undecyloxy)hexyl)amino)octanoate (Compound 1).

139. The composition of claim 137 or 138, wherein the neutral lipid is 1,2 distearoyl-sn-glycero-3-phosphocholine (DSPC).

140. The composition of any one of claims 131-139, wherein the sterol is cholesterol.

141. The composition of any one of claims 131-140, wherein the PEG-modified lipid is 1,2 dimyristoyl-sn-glycerol, methoxypolyethyleneglycol (PEG2000 DMG).

142. The composition of any one of claims 131-141, wherein the lipid nanoparticle comprises 20-60 mol% ionizable cationic lipid, 5-25 mol% neutral lipid, 25-55 mol% sterol, and 0.5-15 mol% PEG-modified lipid.

143. The composition of claim 142, wherein the lipid nanoparticle comprises:
40-50 mol% ionizable lipid, optionally 45-50 mol%, 30-45 mol% sterol, optionally 35-40 mol%, 5-15 mol% helper lipid, optionally 10-12 mol%, 1-5% PEG lipid, optionally 1-3 mol%, or 1.5 to 2.5 mol%.

144. The composition of claim 143, wherein the lipid nanoparticle comprises:
40-50 mol% Compound 1, optionally 45-50 mol%, 30-45 mol% cholesterol, optionally 35-40 mol%, 5-15 mol% DSPC, optionally 10-12 mol%, 1-5% PEG2000DMG, optionally 1-3 mol%, or 1.5 to 2.5 mol%.

145. The mRNA of any one of claims 85-130, wherein the mRNA comprises a chemical modification, optionally 1-methylpseudouridine.

146. The mRNA of any one of claims 85-130, wherein each uridine in the mRNA is a 1-methylpseudouridine.

147. The composition of any one of claims 131-144, wherein the mRNA comprises a chemical modification, optionally 1-methylpseudouridine.

148. The composition of any one of claims 131-144, wherein each uridine in the mRNA is a 1-methylpseudouridine.

149. A method comprising administering to a subject the mRNA of any one of claims 85-130 in an amount effective to induce in the subject a neutralizing antibody response against SARS-CoV-2.

150. A method comprising administering to a subject the mRNA of any one of claims 85-130 in an amount effective to induce in the subject a T cell immune response against SARS-CoV-2.

151. A method comprising administering to a subject the mRNA of any one of claims 85-130 in an amount effective to induce in the subject a neutralizing antibody response and a T cell immune response against SARS-CoV-2.

152. A method comprising administering to a subject the compostion of any one of claims 131-144 in an amount effective to induce in the subject a neutralizing antibody response against SARS-CoV-2.

153. A method comprising administering to a subject the compostion of any one of claims 131-144 in an amount effective to induce in the subject a T cell immune response against SARS-CoV-2.

154. A method comprising administering to a subject the compostion of any one of claims 131-144 in an amount effective to induce in the subject a neutralizing antibody response and a T cell immune response against SARS-CoV-2.