CORONAVIRUS RNA VACCINES

RELATED APPLICATION

This application claims the benefit under 35 U.S.C. § 119(e) of U.S. provisional application number 62/967,006, filed January 28, 2020, U.S. provisional application number 62/971,825, filed February 7, 2020, U.S. provisional application number 63/002,094, filed March 30, 2020, U.S. provisional application number 63/009,005, filed April 13, 2020, and U.S. provisional application number 63/016,175, filed April 27, 2020, each of which are incorporated by reference herein in their 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 (beta-coronaviruses). 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 Wuhan coronavirus, however, have been associated with a high mortality rate death toll. A Severe Acute Respiratory Syndrome Coronavirus 2 (SARSCoV-2) (formerly referred to as a "Wuhan coronavirus," a "2019 novel coronavirus," or a "2019-nCoV") was initially identified from the Chinese city of Wuhan in December 2019 and has rapidly infected hundreds of thousands of people. The pandemic disease that the SARSCoV-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 (also referred to as 2019 nCoV or Wuhan-Hu-1) was deposited in GenBank on January 12, 2020 by investigators from the Chinese CDC in Beijing.

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 immunizing compositions (e.g., RNA vaccines) that comprise an RNA that encodes highly immunogenic antigens capable of eliciting potent neutralizing antibodies responses against coronavirus antigens, such as SARS-CoV-2

antigens. Surprisingly, the protein antigen sequences of this novel coronavirus share less than 80% identity with the protein antigen sequences of Severe Acute Respiratory Syndrome (SARS) coronavirus, and less than 35% identity with the protein antigen sequences of the Middle East Respiratory Syndrome (MERS) coronavirus.

The constructs provided herein, in some embodiments, include: a reversion of the polybasic cleavage site in the native SARS-CoV-2 Spike (S) protein to a single basic cleavage site (e.g., FIG. 1, Variant 7, SEQ ID NO: 23); a deletion of the polybasic ER/Golgi signal sequence (KXHXX-COOH) at the carboxy tail (e.g., FIG. 1, Variant 8, SEQ ID NO: 26); a double proline stabilizing mutation (e.g., FIG. 1, Variants 1-6 and 9, SEQ ID NOs: 5, 8, 11, 14, 17, 20, and 29); a modified protease cleavage site to stabilize the protein (e.g., FIG. 1, Variants 3 and 5, SEQ ID NOs: 11 and 17); a deletion of the cytoplasmic tail (e.g., FIG. 1, Variants 3, 4, and 6, SEQ ID NOs: 11, 14, and 20); and/or a foldon scaffold (e.g., FIG. 1, Variants 3 and 4, SEQ ID NOs: 11 and 14). The structural features disclosed herein include, for instance, abolishment of furin cleavage site by optionally replacing it with a transmembrane region, a foldon grafted to the C-terminal portion of the spike ectodomain, deleted C-terminal intracellular tail (carboxy tail), Thus, the mRNA provided herein, in some embodiments, comprises an open reading frame encoding a variant trimeric spike protein comprising any one or more of a deleted furin cleavage site, additional foldon sequence as C-terminus, deleted carboxy tail or sequence therein and/or 2 proline mutation.

Some aspects of the present disclosure provide a ribonucleic acid (RNA) comprising an open reading frame (ORF) that encodes a coronavirus antigen (e.g., an S protein, a membrane (M) protein, an envelope (E) protein, a nucleocapsid (NC) protein, or a protein of Table 1) capable of inducing an immune response (e.g., a neutralizing antibody response) to SARS-CoV-2, optionally wherein the RNA is formulated in a lipid nanoparticle.

Other aspects of the present disclosure provide a codon-optimized RNA comprising an ORF that comprises a sequence having at least 80% identity to a wild-type RNA encoding a SARS-CoV-2 antigen, optionally wherein the RNA is formulated in a lipid nanoparticle.

Yet other aspects of the present disclosure provide a chemically-modified RNA comprising an ORF that comprises a sequence having at least 80% identity to a wild-type RNA encoding a SARS-CoV-2 antigen, optionally wherein the RNA is formulated in a lipid nanoparticle.

Still other aspects of the present disclosure provide an RNA comprising an ORF that comprises a sequence having at least 80% identity to the sequence of any one of the sequences of Table 1, e.g., SEQ ID NOs: 3, 7, 10, 13, 16, 19, 22, 25, 28, 31, 48, 50, 52, 54, 56, 61, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, or 84. In some embodiments, the RNA comprises an ORF that

comprises a sequence having at least 80% identity to the sequence of SEQ ID NO: 28. In some embodiments, the RNA comprises an ORF that comprises a sequence having at least 80% identity to the sequence of SEQ ID NO: 16. In some embodiments, the RNA comprises an ORF that comprises a sequence having at least 80% identity to the sequence of SEQ ID NO: 19. In some embodiments, the RNA comprises an ORF that comprises a sequence having at least 80% identity to the sequence of SEQ ID NO: 22. In some embodiments, the RNA comprises an ORF that comprises a sequence having at least 80% identity to the sequence of SEQ ID NO: 25.

In some embodiments, the ORF comprises a sequence having at least 85%, at least 90%, at least 95%, or at least 98% identity to the sequence of any one of the sequences of Table 1, e.g., SEQ ID NOs: 3, 7, 10, 13, 16, 19, 22, 25, 28, 31, 48, 50, 52, 54, 56, 61, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, or 84. In some embodiments, the RNA comprises an ORF that comprises a sequence having at least 85%, at least 90%, at least 95%, or at least 98% identity to the sequence of SEQ ID NO: 28. In some embodiments, the RNA comprises an ORF that comprises the sequence of SEQ ID NO: 28. In some embodiments, the RNA comprises an ORF that comprises a sequence having at least 85%, at least 90%, at least 95%, or at least 98% identity to the sequence of SEQ ID NO: 16. In some embodiments, the RNA comprises an ORF that comprises the sequence of SEQ ID NO: 16. In some embodiments, the RNA comprises an ORF that comprises a sequence having at least 85%, at least 90%, at least 95%, or at least 98% identity to the sequence of SEQ ID NO: 19. In some embodiments, the RNA comprises an ORF that comprises the sequence of SEQ ID NO: 19. In some embodiments, the RNA comprises an ORF that comprises a sequence having at least 85%, at least 90%, at least 95%, or at least 98% identity to the sequence of SEQ ID NO: 22. In some embodiments, the RNA comprises an ORF that comprises the sequence of SEQ ID NO: 22. In some embodiments, the RNA comprises an ORF that comprises a sequence having at least 85%, at least 90%, at least 95%, or at least 98% identity to the sequence of SEQ ID NO: 25. In some embodiments, the RNA comprises an ORF that comprises the sequence of SEQ ID NO: 25. In some embodiments, the RNA comprises an ORF that comprises a sequence having at least 85%, at least 90%, at least 95%, or at least 98% identity to the sequence of SEQ ID NO: 106. In some embodiments, the RNA comprises an ORF that comprises the sequence of SEQ ID NO: 106. In some embodiments, mRNAs comprising the ORF are uniformly modified (e.g., fully modified, modified throughout the entire sequence) for a particular modification. For example, an RNA can be uniformly modified with 1-methyl-pseudouridine, such that each U in the sequence is a 1-methyl-pseudouridine.

In some embodiments, the RNA further comprises a 5′ UTR, optionally wherein the 5′ UTR comprises the sequence of SEQ ID NO: 2 or SEQ ID NO: 36.

In some embodiments, the RNA further comprises a 3′ UTR, optionally wherein the 3′ UTR comprises the sequence of SEQ ID NO: 4 or SEQ ID NO: 37.

In some embodiments, the RNA further comprises a 5′ cap analog, optionally a 7mG(5′)ppp(5′)NlmpNp cap. Other cap analogs may be used.

In some embodiments, the RNA further comprises a poly(A) tail, optionally having a length of 50 to 150 nucleotides.

In some embodiments, the ORF encodes a coronavirus antigen. In some embodiments, the coronavirus antigen is a structural protein. In some embodiments, the structural protein is a spike (S) protein. In some embodiments, the S protein is a stabilized prefusion form of an S protein. In some embodiments, the coronavirus antigen comprises a sequence having at least 80% identity to the sequence of any one of the sequences of Table 1, e.g., SEQ ID NOs: 5, 8, 11, 14, 17, 20, 23, 26, 29, 32, 33, 34, 35, 47, 49, 59, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, or 85. In some embodiments, the coronavirus antigen comprises a sequence having at least 80% identity to the sequence of SEQ ID NO: 29. In some embodiments, the coronavirus antigen comprises a sequence having at least 80% identity to the sequence of SEQ ID NO: 17. In some embodiments, the coronavirus antigen comprises a sequence having at least 80% identity to the sequence of SEQ ID NO: 20. In some embodiments, the coronavirus antigen comprises a sequence having at least 80% identity to the sequence of SEQ ID NO: 23. In some embodiments, the coronavirus antigen comprises a sequence having at least 80% identity to the sequence of SEQ ID NO: 26. In some embodiments, the coronavirus antigen comprises a sequence having at least 85%, at least 90%, at least 95%, or at least 98% identity to the sequence of any one of the sequences of Table 1, e.g., SEQ ID NOs: 5, 8, 11, 14, 17, 20, 23, 26, 29, 32, 33, 34, 35, 47, 49, 59, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, or 85. In some embodiments, the coronavirus antigen comprises a sequence having at least 85%, at least 90%, at least 95%, or at least 98% identity to the sequence of SEQ ID NO: 29. In some embodiments, the coronavirus antigen comprises the sequence of SEQ ID NO: 29. In some embodiments, the coronavirus antigen comprises a sequence having at least 85%, at least 90%, at least 95%, or at least 98% identity to the sequence of SEQ ID NO: 17. In some embodiments, the coronavirus antigen comprises the sequence of SEQ ID NO: 17. In some embodiments, the coronavirus antigen comprises a sequence having at least 85%, at least 90%, at least 95%, or at least 98% identity to the sequence of SEQ ID NO: 20. In some embodiments, the coronavirus antigen comprises the sequence of SEQ ID NO: 20. In some embodiments, the coronavirus antigen comprises a sequence having at least 85%, at least 90%, at least 95%, or at least 98% identity to the sequence of SEQ ID NO: 23. In some embodiments, the coronavirus antigen comprises the sequence of SEQ ID NO: 23. In some embodiments, the coronavirus antigen comprises a sequence having at least 85%, at least

90%, at least 95%, or at least 98% identity to the sequence of SEQ ID NO: 26. In some embodiments, the coronavirus antigen comprises the sequence of SEQ ID NO: 26.

In some embodiments, the structural protein is an M protein. In some embodiments, the M protein comprise a sequence having at least 80% identity to the sequence of SEQ ID NO: 81. In some embodiments, the M protein comprises a sequence having at least 85%, at least 90%, at least 95%, or at least 98% identity to the sequence of SEQ ID NO: 81. In some embodiments, the ORF comprises the sequence of SEQ ID NO: 80. In some embodiments, the RNA comprises a sequence having at least 85%, at least 90%, at least 95%, or at least 98% identity to the sequence of SEQ ID NO: 95. In some embodiments, the RNA comprises the sequence of SEQ ID NO: 95.

In some embodiments, the structural protein is an E protein. In some embodiments, the E protein comprises a sequence having at least 80% identity to the sequence of SEQ ID NO: 83. In some embodiments, the E protein comprises a sequence having at least 85%, at least 90%, at least 95%, or at least 98% identity to the sequence of SEQ ID NO: 83. In some embodiments, the ORF comprises the sequence of SEQ ID NO: 82. In some embodiments, the RNA comprises a sequence having at least 85%, at least 90%, at least 95%, or at least 98% identity to the sequence of SEQ ID NO: 96. In some embodiments, the RNA comprises the sequence of SEQ ID NO: 96.

In some embodiments, the structural protein is an NC protein. In some embodiments, the NC protein comprises a sequence having at least 80% identity to the sequence of SEQ ID NO: 85. In some embodiments, the NC protein comprises a sequence having at least 85%, at least 90%, at least 95%, or at least 98% identity to the sequence of SEQ ID NO: 85. In some embodiments, the ORF comprises the sequence of SEQ ID NO: 84. In some embodiments, the RNA comprises a sequence having at least 85%, at least 90%, at least 95%, or at least 98% identity to the sequence of SEQ ID NO: 97. In some embodiments, the RNA comprises the sequence of SEQ ID NO: 97.

In some embodiments, the ORF comprises the sequence of any one of the sequences of Table 1, e.g., any one of SEQ ID NOs: 3, 7, 10, 13, 16, 19, 22, 25, 28, 31, 48, 50, 52, 54, 56, 61, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, or 106. In some embodiments, mRNAs comprising the ORF are uniformly modified (e.g., fully modified, modified throughout the entire sequence) for a particular modification. For example, an RNA can be uniformly modified with 1-methyl-pseudouridine, such that each U in the sequence is a 1-methyl-pseudouridine.

In some embodiments, the RNA comprises a sequence having at least 85%, at least 90%, at least 95%, or at least 98% identity to the sequence of any one of the sequences of Table 1, e.g., any one of SEQ ID NOs: 1, 6, 9, 12, 15, 18, 21, 24, 27, 30, 51, 53, 55, 57, 58, 60, 86-97, or 105. In some embodiments, the RNA comprises a sequence having at least 85%, at least 90%, at least 95%, or at least 98% identity to the sequence of SEQ ID NO: 27. In some embodiments, the

RNA comprises a sequence having at least 85%, at least 90%, at least 95%, or at least 98% identity to the sequence of SEQ ID NO: 105. In some embodiments, the RNA comprises a sequence having at least 85%, at least 90%, at least 95%, or at least 98% identity to the sequence of SEQ ID NO: 15. In some embodiments, the RNA comprises a sequence having at least 85%, at least 90%, at least 95%, or at least 98% identity to the sequence of SEQ ID NO: 18. In some embodiments, the RNA comprises a sequence having at least 85%, at least 90%, at least 95%, or at least 98% identity to the sequence of SEQ ID NO: 21. In some embodiments, the RNA comprises a sequence having at least 85%, at least 90%, at least 95%, or at least 98% identity to the sequence of SEQ ID NO: 24.

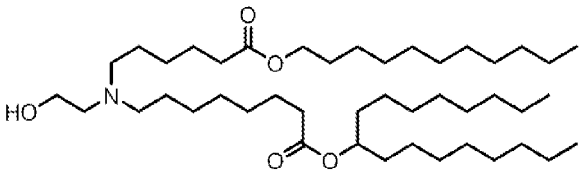
In some embodiments, the RNA comprises the sequence of any one of the sequences of Table 1, e.g., any one of SEQ ID NOs: 1, 6, 9, 12, 15, 18, 21, 24, 27, 30, 51, 53, 55, 57, 58, 60, 86-97, or 105. In some embodiments, the RNA comprises the sequence of SEQ ID NO: 27. In some embodiments, the RNA comprises the sequence of SEQ ID NO: 15. In some embodiments, the RNA comprises the sequence of SEQ ID NO: 18. In some embodiments, the RNA comprises the sequence of SEQ ID NO: 21. In some embodiments, the RNA comprises the sequence of SEQ ID NO: 24. In some embodiments, mRNAs are uniformly modified (e.g., fully modified, modified throughout the entire sequence) for a particular modification. For example, an RNA can be uniformly modified with 1-methyl-pseudouridine, such that each U in the sequence is a 1-methyl-pseudouridine.

In some embodiments, the RNA comprises a chemical modification. In some embodiments, the chemical modification is 1-methylpseudouridine (e.g., fully modified, modified throughout the entire sequence).

Some aspects of the present disclosure provide a method comprising codon optimizing the RNA of any one of the preceding embodiments.

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

In some embodiments, the lipid nanoparticle comprises a PEG-modified lipid, a non-cationic lipid, a sterol, an ionizable cationic lipid, or any combination thereof. In some embodiments, the lipid nanoparticle comprises 0.5-15 mol% (e.g., 0.5-10 mol%, 0.5-5 mol%, or 1-2 mol%) PEG-modified lipid; 5-25 mol% (e.g., 5-20 mol%, or 5-15 mol%) non-cationic (e.g., neutral) lipid; 25-55 mol% (e.g., 30-45 mol% or 35-40 mol%) sterol; and 20-60 mol% (e.g., 40-60 mol%, 40-50 mol% , 45-55 mol%, or 45-50 mol%) ionizable cationic lipid. In some embodiments, the PEG-modified lipid is 1,2 dimyristoyl-sn-glycerol, methoxypolyethyleneglycol (PEG2000 DMG), the non-cationic lipid is 1,2 distearoyl-sn-glycero-3-phosphocholine (DSPC), the sterol is cholesterol; and the ionizable cationic lipid has the structure of Compound 1:



(Compound 1).

Other aspects of the present disclosure provide a composition comprising the RNA of any one of the preceding embodiments and a mixture of lipids. In some embodiments, the mixture of lipids comprises a PEG-modified lipid, a non-cationic lipid, a sterol, an ionizable cationic lipid, or any combination thereof. In some embodiments, the mixture of lipids comprises 0.5-15 mol% (e.g., 0.5-10 mol%, 0.5-5 mol%, or 1-2 mol%) PEG-modified lipid; 5-25 mol% (e.g., 5-20 mol%, or 5-15 mol%) non-cationic (e.g., neutral) lipid; 25-55 mol% (e.g., 30-45 mol% or 35-40 mol%) sterol; and 20-60 mol% (e.g., 40-60 mol%, 40-50 mol%, 45-55 mol%, or 45-50 mol%) ionizable cationic lipid. In some embodiments, the PEG-modified lipid is 1,2 dimyristoyl-sn-glycerol, methoxypolyethyleneglycol (PEG2000 DMG), the non-cationic lipid is 1,2 distearoyl-sn-glycero-3-phosphocholine (DSPC), the sterol is cholesterol; and the ionizable cationic lipid has the structure of Compound 1.

In some embodiments, the mixture of lipids forms lipid nanoparticles. In some embodiments, the RNA is formulated in the lipid nanoparticles. In some embodiments, the lipid nanoparticles are formed first as empty lipid nanoparticles and combined with the mRNA of the vaccine immediately prior to (e.g., within a couple of minutes to an hour of) administration.

Yet other aspects of the present disclosure provide a method comprising administering to a subject the RNA of any one of the preceding embodiments in an amount effective to induce a neutralizing antibody response against a coronavirus in the subject.

Still other aspects of the present disclosure provide a method comprising administering to a subject the composition of any one of the preceding embodiments in an amount effective to induce a neutralizing antibody response and/or a T cell immune response, optionally a $CD4^+$ and/or a $CD8^+$ T cell immune response against a coronavirus in the subject.

In some embodiments, the coronavirus is a SARS-CoV-2.

In some embodiments, the subject is immunocompromised. In some embodiments, the subject has a pulmonary disease. In some embodiments, the subject is 5 years of age or younger, or 65 years of age or older.

In some embodiments, the method comprises administering to the subject at least two doses of the composition.

In some embodiments, detectable levels of the coronavirus antigen are produced in serum of the subject at 1-72 hours post administration of the RNA or composition comprising the RNA.

In some embodiments, a neutralizing antibody titer of at least 100 NU/ml, at least 500 NU/ml, or at least 1000 NU/ml is produced in the serum of the subject at 1-72 hours post administration of the RNA or composition comprising the RNA.

It should be understood that the terms "SARS-CoV-2," "Wuhan coronavirus," "2019 novel coronavirus," and "2019-nCoV" refer to the same recently emerged betacoronavirus now known as SARS-CoV-2 and are used interchangeably herein.

The entire contents of International Application No. PCT/US2016/058327 (Publication No. WO2017/07062) and International Application No. PCT/US2018/022777 (Publication No. WO2018/170347) are incorporated herein by reference.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 shows schematics of various exemplary S protein antigen encoded by the SARS-CoV-2 mRNA of the present disclosure. The top schematic represents a wild-type SARS-CoV-2 protein; the schematics below depict SARS-CoV-2 protein variants, relative to the wild-type.

FIG. 2 shows a graph of 24 hour in vitro expression data for various SARS-CoV-2 protein variants encoded by the SARS-CoV-2 mRNA of the present disclosure.

FIG. 3 shows graph of 24 hour in vitro expression data for various SARS-CoV-2 protein variants encoded by the SARS-CoV-2 mRNA of the present disclosure. Two different amounts of mRNA were tested.

FIGs. 4A-4B show graphs of serum antibody titer measurements following immunization with different doses of the SARS-CoV-2 Variant 9 mRNA vaccine in different strains of mice (FIG. 4A) and at higher doses (FIG. 4B).

FIGs. 5A-5C show graphs of serum antibody titer measurements following immunization with different doses of the SARS-CoV-2 Variant 5 mRNA vaccine (FIG. 5A), compared to the SARS-CoV-2 Variant 9 mRNA vaccine and an mRNA encoding a wild-type SARS-CoV-2 S protein (FIG. 5B). FIG. 5C is a graph comparing the serum antibody titer for seven different SARS-CoV-2 mRNA vaccines and an mRNA encoding wild-type SARS-CoV-2 S protein Sequence.

FIG. 6 shows a graph of a temporal antibody response in mice after immunization with the SARS-CoV-2 Variant 9 mRNA at different doses.

FIG. 7 shows a schematic depicting dosing schedules.

FIGs. 8A-8C show graphs of serum antibody titers in mice two weeks after a priming dose of the SARS-CoV-2 Variant 9 mRNA vaccine and two weeks after a booster dose of the Wuahn-Hu-1 Variant 9 mRNA vaccine in BALB/c mice (FIG. 8A), C57BL/6 mice (FIG. 8B), and C3B6 mice (FIG. 8C). Various vaccine doses were tested.

FIGs. 9A-9E show graphs of serum antibody titers from mice two weeks after a priming dose of the SARS-CoV-2 Variant 5 mRNA vaccine and two weeks after a booster dose of the SARS-CoV-2 Variant 5 mRNA vaccine in BALB/c mice (FIG. 9A) and in C3B6 mice (FIG. 9B) or after a priming dose and booster dose of mRNA encoding wild-type SARS-CoV-2 protein (FIG. 9C). Various vaccine doses were tested. FIGs. 9D-9E show graphs comparing serum antibody titers in BALB/c mice (FIG. 9D) and in C3B6 mice (FIG. 9E) immunized with the SARS-CoV-2 Variant 9 mRNA vaccine, the SARS-CoV-2 Variant 5 mRNA vaccine, or mRNA encoding wild-type SARS-CoV-2 S protein.

FIG. 10 shows a graph comparing the serum antibody titer from mice immunized with one of seven different SARA-CoV-2 mRNA vaccines or an mRNA encoding wild-type SARS-CoV-2 S protein sequence following a booster dose.

FIGs. 11A-11B show graphs of the results of a flow cytometry analysis using the 5653-118 ("118") antibody, which is specific for the N-terminal domain of SARS-CoV-1 S1 subunit, following immunization of mice with the SARS-CoV-2 Variant 9 mRNA vaccine, the SARS-CoV-2 Variant 5 mRNA vaccine, or the SARS-CoV-2 Variant 6 mRNA vaccine. The analysis was performed using lymph node (FIG. 11A) and spleen (FIG. 11B) samples obtained from the mice.

FIGs. 12A-12B show graphs of the results of a flow cytometry analysis using the 5652-109 ("109") antibody, which is specific for the receptor-binding domain of SARS-CoV-1 S protein, following immunization of mice with the SARS-CoV-2 Variant 9 mRNA vaccine, the SARS-CoV-2 Variant 5 mRNA vaccine, or the SARS-CoV-2 Variant 6 mRNA vaccine. The analysis was performed using lymph node (FIG. 12A) and spleen (FIG. 12B) samples obtained from the mice.

FIGs. 13A-13C show graphs of the results of a flow cytometry analysis following transfection with one of six different SARS-CoV-2 mRNA vaccines in vitro. FIG. 13A shows the percentage of antigen-presenting cell-positive (APC+), and FIG. 13B shows the mean fluorescence intensity (MFI). FIG. 13C shows results using a positive control (a SARS antibody).

FIG. 14 shows graphs of the results from a flow cytometry analysis following transfection with the SARS-CoV-2 Variant 9 mRNA vaccine in vitro, using mAb118, mAb109, and SARS mAb103 (positive control). The negative control excluded a primary antibody.

FIG. 15 shows graphs of protein binding between mAb118 or mAb109 and a SARS-CoV-2 antigen at different concentrations.

FIGs. 16A-16B show graphs of binding and neutralizing antibodies in BALB/c mice vaccinated with 1 μg, 0.1 μg or 0.01 μg of the SARS-CoV-2 Variant 9 mRNA vaccine at weeks

0 and 3. FIG. 16A shows S-2P-binding antibodies assessed by ELISA at week 2 (post-prime) and week 5 (post-boost). FIG. 16B shows neutralizing activity assessed at week 5 by a pseudovirus neutralization assay in sera of mice that received 1 μg or 0.1 μg of the SARS-CoV-2 Variant 9 mRNA vaccine.

FIGs. 17A-17C show graphs of data demonstrating that SARS-CoV-2 Variant 9 mRNA vaccine-induced immunity prevents SARS-CoV-2 replication in the lungs of BALB/c mice. BALB/c mice were vaccinated with 1 μg, 0.1 μg or 0.01 μg of the SARS-CoV-2 Variant 9 mRNA vaccine at weeks 0 and 3 and challenged at week 9 with mouse-adapted SARS-CoV-2. FIG. 17A shows viral titers in lung assessed by plaque assay on day 2 post-challenge. FIG. 17B shows viral titers in nasal turbinates assessed by plaque assay on day 2 post-challenge. FIG. 17C shows the change in body weight (as a percentage) over time following infection.

FIGs. 18A-18C show graphs of data demonstrating that SARS-CoV-2 Variant 9 mRNA vaccine-induced immunity prevents SARS-CoV-2 replication in the lungs of BALB/c mice. BALB/c mice were vaccinated with 1 μg, 0.1 μg or 0.01 μg of the SARS-CoV-2 Variant 9 mRNA vaccine at week 0 and challenged at week 7 with mouse-adapted SARS-CoV-2. FIG. 18A shows viral titers in nasal turbinates assessed by plaque assay on day 2 post-challenge. FIG. 18B shows viral titers in lung assessed by plaque assay on day 2 post-challenge. FIG. 18C shows the change in body weight (as a percentage) over time following infection.

FIG. 19 shows the week 0 and 3 immunization schedule used in Example 10.

FIGs. 20A-20C show graphs of data demonstrating that SARS-CoV-2 Variant 9 mRNA vaccine-induced immunity prevents SARS-CoV-2 replication in the lungs of BALB/c mice. BALB/c mice were vaccinated with 10 μg, 1 μg or 0. 1 μg of SARS-CoV-2 Variant 9 at weeks 0 and 4 and challenged at week 7 with mouse-adapted SARS-CoV-2. FIG. 20A shows viral titers in nasal turbinates assessed by plaque assay on day 2 post-challenge. FIG. 20B shows viral titers in lung assessed by plaque assay on day 2 post-challenge. FIG. 20C shows the change in body weight (as a percentage) over time following infection.

FIGs. 21A-21H show graphs of data relating to neutralizing antibody responses following mRNA immunization of BALB/c mice. Sigmoidal curves, taking averages of triplicates at each serum dilution, were generated from relative luciferase units (RLU) readings and 50% (IC_{50}) (FIGs. 21A, 21C, 21E, 21G) and 80% (IC_{80}) (FIGs. 21B, 21D, 21F, 21H) neutralizing activity was calculated considering uninfected cells to represent 100% neutralization and cells transduced with only virus to represent 0% neutralization. Each symbol represents an individual mouse, bars represent geometric mean titers (GMT), and error bars indicate geometric standard deviation (SD). FIGs. 21A-21F show unpaired T-tests used to compare 0.1 μg and 1μg

doses. FIGs. 21G and 21H show groups compared by one-way ANOVA with Kruskal-Wallis multiple comparison test.

FIGs. 22A-22C show graphs of data relating to binding and neutralizing antibody responses following low dose mRNA immunization of BALB/c mice with alternative spike antigen designs. FIG. 22A shows serum endpoint titers. FIG. 22B shows 50% (IC_{50}) neutralizing activity calculated considering uninfected cells representing 100% neutralization and cells transduced with only virus representing 0% neutralization. Each symbol represents an individual mouse, bars represent geometric mean titers (GMT), and error bars indicate geometric standard deviation (SD). In FIGs. 22A and 22B, groups were compared by one-way ANOVA with Kruskal-Wallis multiple comparison test. FIG. 22C shows antibody binding and neutralization titers compared by Spearman correlation.

DETAILED DESCRIPTION

The present disclosure provides compositions (e.g., immunizing/immunogenic compositions such as RNA vaccines in lipid nanoparticles) that elicit potent neutralizing antibodies against coronavirus antigens. In some embodiments, an immunizing composition includes RNA (e.g., messenger RNA (mRNA)) encoding a coronavirus antigen, such as a SARS-CoV-2 antigen in a lipid nanoparticle. In some embodiments, the coronavirus antigen is a structural protein. In some embodiments, the coronavirus antigen is a spike protein, an envelope protein, a nucleocapsid protein, or a membrane protein. In some embodiments, the coronavirus antigen is a stabilized prefusion spike protein. In some embodiments, the mRNA comprises an open reading frame that encodes a variant trimeric spike protein. The trimeric spike protein, for example, may comprise a stabilized prefusion spike protein. In some embodiments, the stabilized prefusion spike protein a double proline (S2P) mutation.

Antigens

Antigens, as used herein, 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 and immunogenic fragments (an immunogenic fragment that induces (or is capable of inducing) an immune response to a (at least one) coronavirus), unless otherwise stated. It should be understood that the term "protein" encompasses 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 provided herein.

The mRNA provided herein, in some embodiments, comprises an open reading frame encoding a variant trimeric spike protein. In some embodiments, the open reading frame encodes a variant trimeric spike protein that comprises a stabilized prefusion spike protein. The stabilized prefusion spike protein, in some embodiments, comprises a double proline (S2P) mutation.

Exemplary sequences of the coronavirus antigens and the RNA (e.g., mRNA) encoding the coronavirus antigens of the compositions of the present disclosure are provided in Table 1.

In some embodiments, a composition comprises an RNA (e.g., mRNA) that encodes a coronavirus antigen that comprises an amino acid sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, or 100% identity to the amino acid sequence of any one of SEQ ID NOs: 5, 8, 11, 14, 17, 20, 23, 26, 29, 32, 33, 34, 35, 47, 49, 59, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, or 85 and optionally a lipid nanoparticle. In some embodiments, a composition comprises an RNA (e.g., mRNA) that encodes a coronavirus antigen that comprises the sequence of SEQ ID NO: 29. In some embodiments, a composition comprises an RNA (e.g., mRNA) that encodes a coronavirus antigen that comprises the sequence of SEQ ID NO: 17. In some embodiments, a composition comprises an RNA (e.g., mRNA) that encodes a coronavirus antigen that comprises the sequence of SEQ ID NO: 20. In some embodiments, a composition comprises an RNA (e.g., mRNA) that encodes a coronavirus antigen that comprises the sequence of SEQ ID NO: 23. In some embodiments, a composition comprises an RNA (e.g., mRNA) that encodes a coronavirus antigen that comprises the sequence of SEQ ID NO: 26.

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

Nucleic Acids

The compositions of the present disclosure comprise a (at least one) RNA having an open reading frame (ORF) encoding a coronavirus antigen (e.g., variant trimeric spike protein, such as a stabilized prefusion spike protein). In some embodiments, the RNA is a messenger RNA (mRNA). In some embodiments, the RNA (e.g., mRNA) further comprises a 5' UTR, 3' UTR, a poly(A) tail and/or a 5' cap analog.

It should also be understood that the coronavirus vaccine 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, 36, 4, or 37); however, other UTR sequences may be used or exchanged for any of the UTR sequences described herein. UTRs may also be omitted from the RNA polynucleotides provided herein.

Nucleic acids comprise a polymer of nucleotides (nucleotide monomers). Thus, nucleic acids are also referred to as polynucleotides. Nucleic acids may be or may include, for example, deoxyribonucleic acids (DNAs), ribonucleic acids (RNAs), 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.

Messenger RNA (mRNA) is any 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 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 RNA (e.g., 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 RNA (e.g., mRNA) sequence complementary to the DNA, where each "T" of the DNA sequence is substituted with "U."

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 RNA polynucleotide of the present disclosure.

In some embodiments, a composition comprises an RNA (e.g., mRNA) that comprises a nucleotide sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, or 100% identity to the nucleotide sequence of any one of SEQ ID NOs: 1, 6, 9, 12, 15, 18, 21, 24, 27, 30, 51, 53, 55, 57, 58, 60, or 86-97.

In some embodiments, a composition comprises an RNA (e.g., mRNA) that comprises an ORF that comprises a nucleotide sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, or 100% identity to the nucleotide sequence of any one of SEQ ID NOs: 3, 7, 10, 13, 16, 19, 22, 25, 28, 31, 48, 50, 52, 54, 56, 61, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, or 84.

Variants

In some embodiments, the compositions of the present disclosure include RNA that encodes a coronavirus antigen variant (e.g., variant trimeric spike protein, such as a stabilized prefusion spike protein). 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 RNA or an RNA ORF that comprises a nucleotide sequence of any one of the sequences provided herein (see, e.g., Sequence Listing and Table 1), 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, in particular 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 RNA (e.g., 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 RNA polynucleotide 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 RNA (e.g., 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 RNA (e.g., 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 RNA (e.g., 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 RNA (e.g., 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 RNA (e.g., 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 RNA (e.g., 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. In some embodiments, the signal peptide may comprise one of the following sequences: MDSKGSSQKGSRLLLLLVVSNLLLPQGVVG (SEQ ID NO: 38), MDWTWILFLVAAATRVHS (SEQ ID NO: 39); METPAQLLFLLLLWLPDTTG (SEQ ID NO: 40); MLGSNSGQRVVFTILLLLVAPAYS (SEQ ID NO: 41); MKCLLYLAFLFIGVNCA (SEQ ID NO: 42); MWLVSLAIVTACAGA (SEQ ID NO: 43); or MFVFLVLLPLVSSQC (SEQ ID NO: 99).

Fusion Proteins

In some embodiments, a composition of the present disclosure includes an RNA (e.g., 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.

Scaffold Moieties

The RNA (e.g., mRNA) vaccines as provided herein, in some embodiments, encode fusion proteins that comprise coronavirus antigens linked to scaffold moieties. In some embodiments, such scaffold moieties impart desired properties to an antigen encoded by a nucleic acid of the disclosure. For example, scaffold proteins 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 a binding partner.

In some embodiments, the scaffold moiety is protein that can self-assemble into protein nanoparticles that are highly symmetric, stable, and structurally organized, with diameters of 10–150 nm, a highly suitable size range for optimal interactions with various cells of the immune

system. In some embodiments, viral proteins or virus-like particles can be used to form stable nanoparticle structures. Examples of such viral proteins are known in the art. For example, in some embodiments, the scaffold moiety is a hepatitis B surface antigen (HBsAg). HBsAg forms spherical particles with an average diameter of ~22 nm and which lacked nucleic acid and hence are non-infectious (Lopez-Sagaseta, J. et al. Computational and Structural Biotechnology Journal 14 (2016) 58–68). In some embodiments, the scaffold moiety is a hepatitis B core antigen (HBcAg) self-assembles into particles of 24–31 nm diameter, which resembled the viral cores obtained from HBV-infected human liver. HBcAg produced in self-assembles into two classes of differently sized nanoparticles of 300 Å and 360 Å diameter, corresponding to 180 or 240 protomers. In some embodiments, the coronavirus antigen is fused to HBsAG or HBcAG to facilitate self-assembly of nanoparticles displaying the coronavirus antigen.

In some embodiments, bacterial protein platforms may be used. Non-limiting examples of these self-assembling proteins include ferritin, lumazine and encapsulin.

Ferritin is a protein whose main function is intracellular iron storage. Ferritin is made of 24 subunits, each composed of a four-alpha-helix bundle, that self-assemble in a quaternary structure with octahedral symmetry (Cho K.J. et al. J Mol Biol. 2009;390:83–98). Several high-resolution structures of ferritin have been determined, confirming that Helicobacter pylori ferritin is made of 24 identical protomers, whereas in animals, there are ferritin light and heavy chains that can assemble alone or combine with different ratios into particles of 24 subunits (Granier T. et al. J Biol Inorg Chem. 2003;8:105–111; Lawson D.M. et al. Nature. 1991;349:541–544). Ferritin self-assembles into nanoparticles with robust thermal and chemical stability. Thus, the ferritin nanoparticle is well-suited to carry and expose antigens.

Lumazine synthase (LS) is also well-suited as a nanoparticle platform for antigen display. LS, which is responsible for the penultimate catalytic step in the biosynthesis of riboflavin, is an enzyme present in a broad variety of organisms, including archaea, bacteria, fungi, plants, and eubacteria (Weber S.E. Flavins and Flavoproteins. Methods and Protocols, Series: Methods in Molecular Biology. 2014). The LS monomer is 150 amino acids long, and consists of beta-sheets along with tandem alpha-helices flanking its sides. A number of different quaternary structures have been reported for LS, illustrating its morphological versatility: from homopentamers up to symmetrical assemblies of 12 pentamers forming capsids of 150 Å diameter. Even LS cages of more than 100 subunits have been described (Zhang X. et al. J Mol Biol. 2006;362:753–770).

Encapsulin, a novel protein cage nanoparticle isolated from thermophile Thermotoga maritima, may also be used as a platform to present antigens on the surface of self-assembling nanoparticles. 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 (Rahmanpour R. et al. FEBS J. 2013, 280: 2097-2104).

In some embodiments, an RNA of the present disclosure encodes a coronavirus antigen (e.g., SARS-CoV-2 S protein) fused to a foldon domain. The foldon domain may be, for example, obtained from bacteriophage T4 fibritin (see, e.g., Tao Y, et al. Structure. 1997 Jun 15; 5(6):789-98).

Linkers and Cleavable Peptides

In some embodiments, the mRNAs of the disclosure encode more than one polypeptide, referred to herein as fusion proteins. In some embodiments, the mRNA further encodes a linker located between at least one or each domain of the fusion protein. The linker can be, for example, a cleavable linker or protease-sensitive linker. In some embodiments, the linker is selected from the group consisting of F2A linker, P2A linker, T2A linker, E2A linker, and combinations thereof. This family of self-cleaving peptide linkers, referred to as 2A peptides, has been described in the art (see for example, Kim, J.H. et al. (2011) PLoS ONE 6:e18556). In some embodiments, the linker is an F2A linker. In some embodiments, the linker is a GGGS (SEQ ID NO: 98) linker. In some embodiments, the fusion protein contains three domains with intervening linkers, having the structure: domain-linker-domain-linker-domain.

Cleavable linkers known in the art may be used in connection with the disclosure. Exemplary such linkers include: F2A linkers, T2A linkers, P2A linkers, E2A linkers (See, e.g., WO2017/127750). The skilled artisan will appreciate that other art-recognized linkers may be suitable for use in the constructs of the disclosure (e.g., encoded by the nucleic acids of the disclosure). The skilled artisan will likewise appreciate that other polycistronic constructs (mRNA encoding more than one antigen/polypeptide separately within the same molecule) may be suitable for use as provided herein.

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 RNA 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 RNA (e.g., 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 RNA 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, a 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, a 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, a mRNA of the disclosure comprises pseudouridine (ψ) substitutions at one or more or all uridine positions of the nucleic acid.

In some embodiments, a 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, a 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: 44), 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 (US2014/0206753, WO2013/185069), the sequence GGGAUCCUACC (SEQ ID NO: 45) (WO2014/144196) 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/2015/101414, WO2015/101415, WO/2015/062738, WO2015/024667, WO2015/024667; 5′ UTR element derived from ribosomal protein Large 32 (L32) gene (WO/2015/101414, WO2015/101415, WO/2015/062738), 5′ UTR element derived from the 5′ UTR of an hydroxysteroid (17-β) dehydrogenase 4 gene (HSD17B4) (WO2015/024667), or a 5′ UTR element derived from the 5′ UTR of ATP5A1 (WO2015/024667) 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: 2 and SEQ ID NO: 36.

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: 46) 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, US2011/0086907). 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). In addition a2-globin, a1-globin, UTRs and mutants thereof are also known in the art (WO2015/101415, WO2015/024667). 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, US2014/0206753, 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 (WO2014/144196) is used. In some embodiments, 3′ UTRs of human and mouse ribosomal protein are used. Other examples include rps9 3′UTR (WO2015/101414), FIG4 (WO2015/101415), and human albumin 7 (WO2015/101415).

In some embodiments, a 3' UTR of the present disclosure comprises a sequence selected from SEQ ID NO: 4 and SEQ ID NO: 37.

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.2010/0293625 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 2010/0129877, 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. 2009/0226470, 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 a RNA polynucleotide, 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 a 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 RNA (e.g., 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 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.

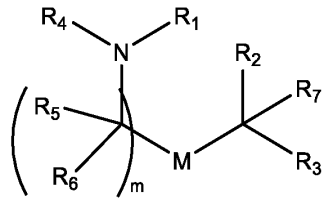
In some embodiments, the lipid nanoparticle comprises 5-25 mol% non-cationic lipid. For example, the lipid nanoparticle may comprise 5-20 mol%, 5-15 mol%, 5-10 mol%, 10-25 mol%, 10-20 mol%, 10-25 mol%, 15-25 mol%, 15-20 mol%, or 20-25 mol% non-cationic lipid. In some embodiments, the lipid nanoparticle 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, $-S(O)_2R$, $-S(O)_2N(R)_2$, 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.

In some embodiments, a subset of compounds of Formula (I) includes those in which when R_4 is $-(CH_2)_nQ$, $-(CH_2)_nCHQR$, –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, $-(CH_2)_nQ$, $-(CH_2)_nCHQR$, -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, $-C(O)N(R)_2$, $-N(R)C(O)R$, $-N(R)S(O)_2R$, $-N(R)C(O)N(R)_2$, $-N(R)C(S)N(R)_2$, $-CRN(R)_2C(O)OR$, $-N(R)R_8$, $-O(CH_2)_nOR$, $-N(R)C(=NR_9)N(R)_2$, $-N(R)C(=CHR_9)N(R)_2$, $-OC(O)N(R)_2$, $-N(R)C(O)OR$, $-N(OR)C(O)R$, $-N(OR)S(O)_2R$, $-N(OR)C(O)OR$, $-N(OR)C(O)N(R)_2$, $-N(OR)C(S)N(R)_2$, $-N(OR)C(=NR_9)N(R)_2$, $-N(OR)C(=CHR_9)N(R)_2$, $-C(=NR_9)N(R)_2$, $-C(=NR_9)R$, -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, -$S(O)_2N(R)_2$, 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 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)$S(O)_2$R, -$N(R)C(O)N(R)_2$, -$N(R)C(S)N(R)_2$, -$CRN(R)_2C(O)OR$, -$N(R)R_8$, -$O(CH_2)_nOR$, -$N(R)C(=NR_9)N(R)_2$, -$N(R)C(=CHR_9)N(R)_2$, -$OC(O)N(R)_2$, -N(R)C(O)OR, -N(OR)C(O)R, -N(OR)$S(O)_2$R, -N(OR)C(O)OR, -$N(OR)C(O)N(R)_2$, -$N(OR)C(S)N(R)_2$, -$N(OR)C(=NR_9)N(R)_2$, -$N(OR)C(=CHR_9)N(R)_2$, -$C(=NR_9)$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;

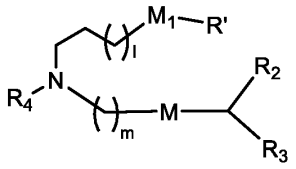
each Y is independently a C_{3-6} carbocycle;

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.

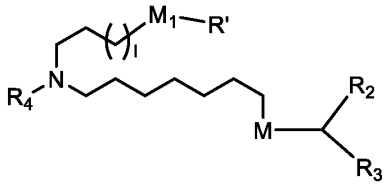
In some embodiments, a subset of compounds of Formula (I) includes those of Formula (IA):



(IA),

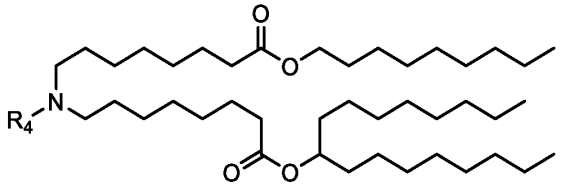
or a salt or isomer thereof, wherein l is selected from 1, 2, 3, 4, and 5; m is selected from 5, 6, 7, 8, and 9; M_1 is a bond or M'; R_4 is unsubstituted C_{1-3} alkyl, or -$(CH_2)_nQ$, in which Q is OH, -NHC(S)N$(R)_2$, -NHC(O)N$(R)_2$, -N(R)C(O)R, -N(R)S$(O)_2$R, -N(R)R_8, -NHC(=NR_9)N$(R)_2$, -NHC(=CHR_9)N$(R)_2$, -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):

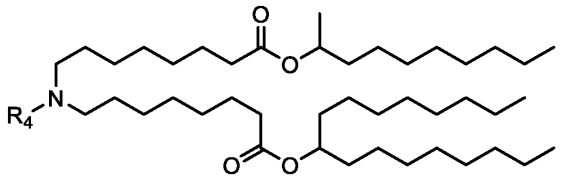


(II) or a salt or isomer thereof, wherein l is selected from 1, 2, 3, 4, and 5; M_1 is a bond or M'; R_4 is unsubstituted C_{1-3} alkyl, or $-(CH_2)_nQ$, in which n is 2, 3, or 4, and Q is OH, $-NHC(S)N(R)_2$, $-NHC(O)N(R)_2$, $-N(R)C(O)R$, $-N(R)S(O)_2R$, $-N(R)R_8$, $-NHC(=NR_9)N(R)_2$, $-NHC(=CHR_9)N(R)_2$, $-OC(O)N(R)_2$, $-N(R)C(O)OR$, 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.

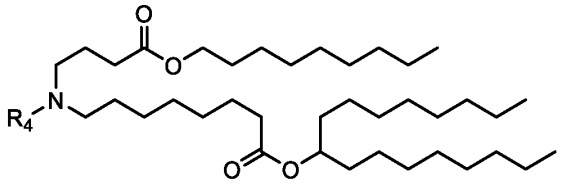
In some embodiments, a subset of compounds of Formula (I) includes those of Formula (IIa), (IIb), (IIc), or (IIe):



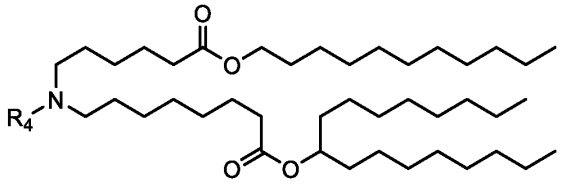
(IIa),



(IIb),



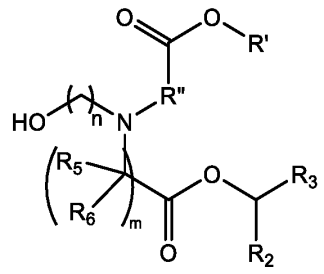
(IIc), or



(IIe),

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

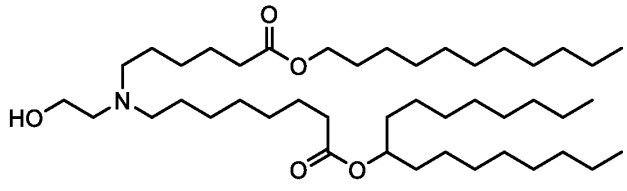
In some embodiments, a subset of compounds of Formula (I) includes those of Formula (IId):



(IId),

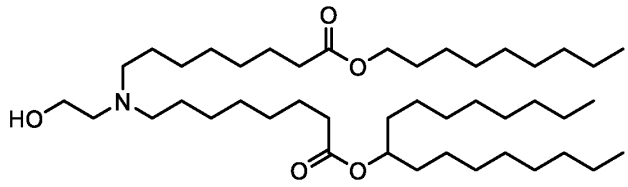
or a salt or isomer thereof, wherein n is 2, 3, or 4; and m, R', R", and R_2 through R_6 are as described herein. For example, each of R_2 and R_3 may be independently selected from the group consisting of C_{5-14} alkyl and C_{5-14} alkenyl.

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), l,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% 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, 36, 37, 38, 39, or 40 mol% cholesterol.

In some embodiments, the lipid nanoparticle comprises 1 – 2 mol% DMG-PEG. For example, the lipid nanoparticle may comprise 1, 1.5, or 2 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, a LNP of the disclosure comprises an N:P ratio of from about 2:1 to about 30:1.

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

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

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

In some embodiments, a 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 RNA or multiple RNAs 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 RNA (e.g., 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 RNA 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 RNA polynucleotides 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 RNA polynucleotides 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, an immunizing composition may comprise other components including, but not limited to, adjuvants.

In some embodiments, an immunizing 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 substances, 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, an immunizing 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, RNA polynucleotides (e.g., mRNA polynucleotides) 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 polynucleotide) 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 RNA 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 immunizing 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, immunizing compositions are used to provide prophylactic protection from coronavirus infection. In some embodiments, immunizing compositions are used to treat a coronavirus infection. In some embodiments, embodiments, immunizing 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, an immunizing 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 an immunizing composition (e.g., an RNA vaccine) 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 an immunizing compositions 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 an immunizing composition comprising a RNA (e.g., 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 an immunizing composition (e.g., an RNA vaccine) comprising a RNA polynucleotide 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 immunizing 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 an immunizing 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 an immunizing 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 an immunizing 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 an immunizing 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 an immunizing 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 an immunizing composition (e.g., an RNA vaccine) comprising an RNA 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 an immunizing 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 RNA 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.

An immunizing composition (e.g., an RNA vaccine) 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 75 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 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 immunizing 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 an immunizing 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.

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 an immunizing 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 n 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 an immunizing 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 an immunizing composition (e.g., RNA vaccine) to be effective is measured in a murine model. For example, an immunizing 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, an immunizing 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 an immunizing 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 an immunizing 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 immunizing composition (e.g., RNA vaccine) is at least 60% relative to unvaccinated control subjects. For example, efficacy of the immunizing 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 an immunizing 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 an immunizing 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 an immunizing 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 an immunizing 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 amount 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 an immunizing 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.

EXAMPLES

Example 1: nCoV In Vitro Expression - DNA

The constructs tested in this experiment were Norwood's DNA co-transfected with a T7 polymerase plasmid to transactivate the promoter on the 2019-nCoV plasmid from Norwood. SARS was used a positive control DNA. The assay conditions were as follows:

DNA constructs:	SARS-CoV-2 Variants 6-10
Cell type:	HEK293T Cells
Plate format:	12-well @ 600,000 cells/well
DNA per well:	2.5 µg/well (construct: T7 =1:1)
Incubation time:	24, 72 hour
Extracellular staining:	single color
Instrument:	LSR Fortessa
ACE2-FLAG, His:	200 µg stock, 10 µg/ml FACS concentration
Anti-FLAG-FITC:	1 mg, 5 µg /ml FACS concentration

Example 2: nCoV In Vitro Expression – mRNA

mRNA of the constructs in Example 1 were tested. The assay conditions were as follows:

mRNA constructs:	SARS-CoV-2 Variants 6-10
Cell type:	HEK293T Cells
Plate format:	24-well @ 300,000 cells/well
mRNA per well:	0.5 µg, 0.1 µg/well
Incubation time:	24, 48 hour
Extracellular staining:	single color
Instrument:	LSR Fortessa
ACE2-FLAG, His:	200 µg stock, 10 µg/ml FACS concentration
Anti-FLAG-FITC:	1 mg, 5 µg /ml FACS concentration

Among all the constructs, SARS-CoV-2 Variant 5 showed best expression as compared with others at low dose. See Figures 2 and 3.

Example 3: Immunogenicity Study

The instant study is designed to test the immunogenicity in mice and/or rabbits of the candidate coronavirus vaccines comprising an mRNA of **Table 1** encoding a coronavirus antigen (e.g., the spike (S) protein, the S1 subunit (S1) of the spike protein, or the S2 subunit (S2) of the spike protein), such as a SARS-CoV-2 antigen.

Animals are vaccinated on week 0 and 3 via intravenous (IV), intramuscular (IM), or intradermal (ID) routes. As controls, one group remains unvaccinated and one is administered inactivated coronavirus. Serum is collected from each animal on weeks 1, 3 (pre-dose) and 5. Individual bleeds are tested for anti-S, anti-S1 or anti-S2 activity via a virus neutralization assay from all three time points, and pooled samples from week 5 only are tested by Western blot using inactivated coronavirus.

In experiments where a lipid nanoparticle (LNP) formulation is used, the formulation may include 0.5-15% PEG-modified lipid; 5-25% non-cationic lipid; 25-55% sterol; and 20-60% ionizable cationic lipid. The PEG-modified lipid may be 1,2 dimyristoyl-sn-glycerol, methoxypolyethyleneglycol (PEG2000 DMG), the non-cationic lipid may be 1,2 distearoyl-sn-glycero-3-phosphocholine (DSPC), the sterol may be cholesterol; and the ionizable cationic lipid may have the structure of Compound 1, for example.

Example 4: Coronavirus Challenge

The instant study is designed to test the efficacy in mice and/or rabbits of candidate coronavirus vaccines comprising an mRNA of **Table 1** encoding a coronavirus antigen (e.g., the spike (S) protein, the S1 subunit (S1) of the spike protein, or the S2 subunit (S2) of the spike protein), such as a SARS-CoV-2 antigen, against a lethal challenge with a coronavirus. Animals are challenged with a lethal dose (10xLD90; ~100 plaque-forming units; PFU) of coronavirus.

The animals used are 6-8 week old female animals in groups of 10. Animals are vaccinated on weeks 0 and 3 via an IM, ID or IV route of administration. Candidate vaccines are chemically modified or unmodified. Animal serum is tested for microneutralization (see Example 14). Animals are then challenged with ~1 LD90 of coronavirus on week 7 via an IN, IM, ID or IV route of administration. Endpoint is day 13 post infection, death or euthanasia. Animals displaying severe illness as determined by >30% weight loss, extreme lethargy or paralysis are euthanized. Body temperature and weight are assessed and recorded daily.

Example 5 – Immunogenicity of SARS-CoV-2 Variant 9 in Mice (one dose)

C3B6, C57/BL6, and BALB/c mice were immunized with various doses of the SARS-CoV-2 Variant 9 mRNA vaccine ("Variant 9") in 50 µL of 1X PBS injected intramuscularly into the right hind leg. Two weeks post-immunization, sera were collected and subjected to ELISA to assess antibody binding to SARS-CoV-2 stabilized prefusion spike protein (SARS-CoV-2 pre-S).

The data for Variant 9 is shown in FIGs. 4A-4B. There were no significant differences between the mouse strains. As shown in FIG. 4A, the C3B6 mice that received 1 µg of Variant 9 had significantly higher antibody responses than the C3B6 mice that received 0.1 µg or 0.01 µg doses (p-value < 0.05). FIG. 4B shows that BALB/c mice that received 10 µg of Variant 9 had significantly higher antibody responses than BALB/c mice that received the 1 µg dose (p-value < 0.05) or the 0.1 µg dose (p-value < 0.0001).

Other mRNA candidates were tested in the manner described above, as shown in FIGs. 5A-5C. FIG. 5A demonstrates that SARS-CoV-2 Variant 5 mRNA vaccine ("Variant 5") elicited similar antibody responses in C3B6 and BALB/c mice following administration of one dose. BALB/c mice that received 1 µg of Variant 9 or Variant 5 had significantly higher antibody responses as compared to BALB/c mice that received 0.1 µg or 0.01 µg doses (p-value < 0.05) (FIG. 5B). At 0.1 µg, Variant 9 elicited similar responses to various other SARS-CoV-2 vaccine antigens delivered by mRNA. Also, at the 0.1 µg dose, SARS-CoV-2 Variant 8 mRNA vaccine ("Variant 8") and SARS-CoV-2 Variant 6 mRNA vaccine ("Variant 6") elicited significantly higher antibody responses than soluble spike protein (S) sequences (* = p-value < 0.05; ** = p-value < 0.01).

Temporal Analysis

BALB/c mice were immunized with various doses of Variant 9 manufactured by the clinically-representative process in 50 µL of 1X PBS intramuscularly into the right hind leg. Every two weeks, post-prime sera were collected and subjected to ELISA to assess antibody binding to SARS-CoV-2 stabilized prefusion spike protein (SARS-CoV-2 pre-S). The results are shown in FIG. 6. Each symbol represents the geometric mean titer (GMT) and error bars indicate the geometric standard deviation (SD). Two-way ANOVA was used to compare antibody responses over time and each dose. The 10 µg dose of Variant 9 was found to elicit a significantly higher antibody response than all other doses (p-value < 0.0001) and significantly decays over 4 weeks after prime (p-value < 0.001).

Example 6 - Immunogenicity of Variant 9 in Mice (two doses)

Comparison of Mouse Strains

Mice (BALB/c, C57BL/6, and C3B6) were immunized with various doses of Variant 9 in 50 µL of 1X PBS intramuscularly into the right hind leg at weeks 0 and 3 (FIG. 7). Two weeks post-prime and post-boost (e.g. weeks 2 and 5), sera were collected and subjected to ELISA to assess antibody binding to SARS-CoV-2 stabilized prefusion spike protein (SARS-CoV-2 pre-S).

The results are shown in FIGs. 8A-8C. Each symbol represents an individual mouse, bars represent geometric mean titers (GMT), and error bars indicate the geometric standard deviation (SD). Two-way ANOVA was used to compare post-prime and post-boost responses. At the 1 µg dose, the BALB/c (FIG. 8A) and C57/BL6 (FIG. 8B) mice immunized with Variant 9 had significantly higher antibody responses following boost (p-value < 0.0001).

Comparison of SARS-CoV-2 mRNA Vaccine Constructs

Mice (BALB/c and C3B6) were immunized with various doses of Variant 9, Variant 5, or SARS-CoV-2 wild-type spike protein mRNA in 50 µL of 1X PBS intramuscularly into the right hind leg at weeks 0 and 3 (FIG. 7). Two weeks post-prime and post-boost (e.g. weeks 2 and 5), sera were collected and subjected to ELISA to assess antibody binding to SARS-CoV-2 stabilized prefusion spike protein (SARS-CoV-2 pre-S).

The results are shown in FIGs. 9A-9E. Each symbol represents an individual mouse, bars represent geometric mean titers (GMT), and error bars indicate the geometric standard deviation (SD). Two-way ANOVA was used to compare post-prime and post-boost responses. At the 1 µg dose, mice immunized with Variant 5 and the spike wild-type sequence (S WT) had significantly higher antibody responses post-boost (p-value < 0.0001) (FIGs. 9A-9C). Also, at the 1 µg dose, the BALB/c mice immunized with Variant 9 had significantly higher antibody responses than mice immunized with the GMP backup sequence (p-value < 0.01) and the S WT (p-value < 0.05) (FIG. 9D). There was no significant difference between antibody responses elicited by either construct in the C3B6 mice (FIG. 9E). For all three sequences tested, there was a significant temporal response (e.g., post-prime dose vs. post-boost dose).

Further research sequences were assayed. BALB/c mice were immunized with various doses of mRNA encoding Variant 9 or other research sequences in 50µL of 1X PBS intramuscularly into the right hind leg at weeks 0 and 3 (FIG. 7). Two weeks post-boost (e.g. week 5), sera were collected and subjected to ELISA to assess antibody binding to SARS-CoV-2 stabilized prefusion spike protein (SARS-CoV-2 pre-S).

The results are shown in FIG. 10. Each symbol represents an individual mouse, bars represent geometric mean titers (GMT), and error bars indicate the geometric standard deviation

(SD). A one-way ANOVA was used to compare all immunogens. Mice immunized with Variant 8, Variant 7, and Variant 6 had significantly higher antibody titers than mice immunized with Variant 9 and S WT (* = p-value < 0.05; ** = p-value < 0.01; **** = p-value < 0.0001).

Example 7 – *In vivo* Expression of SARS-CoV-2 mRNA Vaccine Constructs

Female BALB/c mice, 6-8 weeks of age, were administered either 2 μg or 10 μg of a COVID-19 construct or Tris buffer (as a control) intramuscularly in each hind leg. The constructs comprise Variant 9 in cationic lipid nanoparticles, 10.7 mM sodium acetate, 8.7% sucrose, 20 mM Tris (pH 7.5). Three constructs were tested: Variant 9, Variant 5, and Variant 6. The constructs were stored at -70 ℃ (Variant 9) or -20 ℃ (other constructs). One day later, spleens and lymph nodes were collected to detect protein expression using flow cytometry.

FIGs. 11A-11B show the results using the 5653-118 ("118") antibody, which is specific for the N-terminal domain of SARS-CoV-1 S1 subunit. There was good expression from all the constructs tested, and a dose-dependent drop in expression was observed. In the lymph node (FIG. 11A) and in the spleen (FIG. 11B), Variant 5 showed significantly higher expression (α = 0.05) than either of the other constructs. At the lower dose (2 μg), Variant 9 had significantly higher expression (α = 0.05) than Variant 6.

FIGs. 12A-12B shows the results using the 5652-109 ("109") antibody, which is specific for the receptor-binding domain of SARS-CoV-1 S protein. There was good expression from all the constructs tested, and a dose-dependent drop in expression was observed. There was no significant difference between Variant 9 and Variant 5 at the 10 μg dose in the lymph node or the spleen. At the 2 μg dose, Variant 9 had significantly higher expression (α = 0.05) than the other two constructs in the lymph node (FIG. 12A) and in the spleen (FIG. 12B).

Example 8 – *In vitro* Expression of SARS-CoV-2 mRNA Vaccine Constructs

Six SARS-CoV-2 mRNA vaccine constructs were tested *in vitro*. HEK293t cells were plated (30,000 cells/well) on a 96-well plate. 200 ng of the construct was added to each well and the plates were incubated for 24 hours. Then, the cells were stained with the "118" antibody (at dilutions of 1:100, 1:300, or 1:600), the "109" antibody (at dilutions of 1:100, 1:300, or 1:600), or SARS-103 (binds RBD from SARS-CoV-1 at a dilution of 1:100; as a control). Staining was then performed with anti-human FC AL647, at a dilution of 1:500, and the samples were read using the LSR Fortessa. The results are shown in FIGs. 13A-13C. The results of Variant 9 are provided in FIG. 14.

Example 9 – *In vitro* Potency Assay Development

An assay was developed to examine the potency of different constructs. Two antibodies, 118 (specific for the N-terminal domain of SARS-CoV-1 S1 subunit) and 109 (specific for the receptor-binding domain of SARS-CoV-1 S protein) were tested. As shown in FIG. 15, only the 118 antibody bound SARS-CoV-2 antigen at different concentrations and doses.

Example 10 – SARS-CoV-2 Variant 9 mRNA Vaccine Mouse Immunogenicity Study

Studies were initiated to evaluate immunogenicity and efficacy of low doses of SARS-CoV-2 Variant 9 mRNA vaccine ("Variant 9") in BALB/c mice. BALB/c mice were vaccinated with 1 μg, 0.1 μg or 0.01 μg of Variant 9 at weeks 0 and 3. Binding antibodies to the stabilized S-2P protein were quantified at weeks 2 and 5. Two weeks after a single dose, there were substantial levels of binding antibodies to the S-2P protein measured by ELISA in mice that received 1 μg of Variant 9 (Fig 16A). A second dose of Variant 9 significantly increased the level of binding antibodies in mice receiving 1 μg or 0.1 μg of Variant 9 (FIG. 16A). Neutralizing activity against SARS-CoV-2 was assessed using a pseudotyped lentivirus reporter neutralization assay. Variant 9 elicited significant neutralizing activity at the 1 μg dose compared to mice that received 0.1 μg of Variant 9 (FIG. 16B).

BALB/c mice were immunized with 1 or 0.1 μg of Variant 9, Variant 5, or wild type (WT) without the 2 proline mutation at weeks 0 and 3. Two weeks post-boost, sera were collected and analyzed by pseudotyped lentivirus reporter neutralization assays against homologous SARS-CoV-2. Sigmoidal curves, taking averages of triplicates at each serum dilution, were generated from relative luciferase units (RLU) readings, and 50% (IC_{50}) (FIGs. 21A, 21C, 21E, 21G) and 80% (IC_{80}) (FIGs. 21B, 21D, 21F, 21H) neutralizing activity was calculated considering uninfected cells to represent 100% neutralization and cells transduced with only virus to represent 0% neutralization. As shown FIGs. 21A-21F, mice immunized with 1 μg of SARS-CoV-2 S mRNA had significantly higher neutralizing antibody responses than mice immunized with 0.1 μg (*** = p-value < 0.001, **** = p-value < 0.0001). Further, mice immunized with 0.1 μg dose did not have detectable neutralizing activity, suggesting sub-protective antibody levels. Also, stabilized SARS-CoV-2 S with 2P mutation (Variant 5 and Variant 9) induced more potent IC_{50} neutralizing activity than WT S (*p-value < 0.05). Inclusion of the native S1/S2 furin cleavage site or replacing the furin cleavage site with a GS linker to produce a single-chain construct did not have significant effect on immunogenicity (FIGs. 21G, 21H).

BALB/c mice immunized with 1 μg, 0.1 μg, or 0.01 μg of Variant 9 at weeks 0 and 3 in 50 μL of 1X PBS intramuscularly into the right hind leg and challenged at week 9 intranasally with 1 x 10^5 PFU of a mouse-adapted SARS-CoV-2 which contains two targeted amino acid

changes in the receptor-binding domain to remove clashes with the mouse ACE-2 receptor (see FIG. 19 for schedule). On day 2 post-challenge, mouse lungs and noses were homogenized and assess for viral load by plaque assays. The plaque-forming units in one lobe of lung (FIG. 17A) and in nasal turbinates (FIG. 17B) show the 1 μg dose group is fully protected with 60-fold reduction in titer compared to the control group. In contrast, unimmunized challenged mice had viral loads of about 10^6 PFU per lung lobe. A dose effect was observed: the 0.1 μg Variant 9 dose reduced lung viral load by approximately 2 logs and the 0.01 μg Variant 9 dose reduced lung viral load by approximately 0.5 log.

In a further study, mice were vaccinated with 10 μg, μg, or 0.1 μg of Variant 9 in 50 μL of 1X PBS intramuscularly into the right hind leg one time (week 0) and were challenged intranasally at week 7 with $1x10^5$ PFU of a mouse-adapted SARS-CoV-2 which contains two targeted amino acid changes in the receptor-binding domain to remove clashes with the mouse ACE-2 receptor. On day 2 post-challenge, mouse lungs (FIG. 18A) and noses (FIG. 18B) were homogenized and assessed for viral load by plaque assay. As seen in FIG. 18A, the 10 μg dose and the 1 μg dose groups are fully protected from viral replication in the lung following challenge, with a 60-fold reduction in titer compared to the control group. Percent body weight is shown in FIG. 18C.

In an additional study, mice were vaccinated with 1 μg, 0.1 μg, or 0.01 μg of Variant 9 at weeks 0 and 4 and challenged at week 7 with a mouse-adapted SARS-CoV-2 which contains two targeted amino acid changes in the receptor-binding domain to remove clashes with the mouse ACE-2 receptor. Plaque-forming units in one lobe of lung (FIG. 20A) and in nasal turbinates (FIG. 20B) and at day two post-challenge show that the 1 μg dose group and the 0.1 μg dose group are fully protected, with approximately a 60-fold reduction in titer compared to the control group. Percent body weight is shown in FIG. 20C.

Example 11 - SARS-CoV-2 Variant 9 mRNA Vaccine Compare to Alternative Sequences

This Example provides data relating to binding and neutralizing antibody responses following low dose mRNA immunization with alternative spike antigen designs. BALB/c mice were immunized with 0.1 μg of mRNA encoding different SARS-CoV-2 S-2P variants. Mice were immunized twice at weeks 0 and 3. Two weeks post-boost, sera were collected and analyzed by fold-on competed ELISA against homologous SARS-CoV-2 stabilized spike (FIG. 22A) and pseudotyped lentivirus reporter neutralization assays (FIG. 22B). FIG. 22A shows serum endpoint binding titers, found by taking averages of duplicates of each serum dilution, and calculated as 4-fold above background optical density. Further, sigmoidal curves, taking averages of triplicates at each serum dilution, were generated from relative luciferase units (RLU) readings

an 50% (IC_{50}) neutralizing activity was calculated considering uninfected cells representing 100% neutralization and cells transduced with only virus representing 0% neutralization (FIG. 22B). In addition, antibody binding and neutralization titers were compared by Spearman correlation (FIG. 22C). It was found that the mRNA encoding sequences containing cytoplasmic tail mutations elicited the most potent antibody responses. Additionally, there was a strong correlation between binding antibody titers and neutralizing antibody titers, where applicable.

Methods

SARS-CoV-2 ELISA

To measure antibody binding, an ELISA was performed. SARS-CoV-2 pre-S was coated onto 96-well Nunc MaxiSorp™ flat-bottom plates (ThermoFisher, catalog #: 44-2401-21) in 100 µL of 1X PBS for 16 hours at 4 °C. Plates were washed 3 times with 250 µL PBS-Tween (PBST) (Medicago AB, catalog #: 09-9410-100). To prevent non-specific binding, plates were blocked with 200 µL PBST supplemented with 5% nonfat skim milk (BD Difco™, catalog #: 232100) (blocking buffer) for 1 hour at room temperature (RT). Plates were washed 3 times with 250 µL PBST. Sera were serially diluted (1:100, 4-fold, 8 times) in 100 µL blocking buffer and allowed to bind to antigen for 1 hour at RT, in duplicate. Plates were washed 3 times with 250 µL PBST. 100 mL goat anti-mouse IgG (H+L) cross-adsorbed secondary antibody conjugated to HRP (ThermoFisher, catolog #: G-21040) diluted in blocking buffer was added for 1 hour at RT. Plates were washed 3 times with 250 µL PBST. The enzyme-linked reaction was developed for 10 minutes with 100 µL KPL SureBlue 1-component TMB microwell peroxidase substrate (Sure Blue, catalog #: 5120-0077) and stopped with 100 µL 1N sulfuric acid (ThermoFisher, catolog #: SA 212-1). Spectramax Paradigm (Molecular Devices) was used to detect OD_{450}. Sera endpoint titers were calculated as 4-fold above non-specific secondary antibody binding to antigen.

Additional Embodiments

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

2. A chemically modified ribonucleic acid (RNA) comprising an open reading frame (ORF) that comprises a sequence having at least 80% identity to a wild-type RNA encoding a SARS-CoV-2 antigen, optionally wherein the RNA is formulated in a lipid nanoparticle.

3. A codon-optimized ribonucleic acid (RNA) comprising an open reading frame (ORF) that comprises a sequence having at least 80% identity to a wild-type RNA encoding a SARS-CoV-2 antigen, optionally wherein the RNA is formulated in a lipid nanoparticle.

4. The RNA of paragraph 2 or 3, wherein the SARS-CoV-2 antigen encoded by the wild-type RNA comprises the sequence of SEQ ID NO: 31.

5. A ribonucleic acid (RNA) comprising an open reading frame (ORF) that comprises a sequence having at least 80% identity to the sequence of any one of SEQ ID NOs: 3, 7, 10, 13, 16, 19, 22, 25, 28, 31, 48, 50, 52, 54, 56, 61, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, or 84.

6. The RNA of paragraph 5, wherein the ORF comprises a sequence having at least 85%, at least 90%, at least 95%, or at least 98% identity to the sequence of any one of SEQ ID NOs: 3, 7, 10, 13, 16, 19, 22, 25, 28, 31, 48, 50, 52, 54, 56, 61, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, or 84.

7. The RNA of paragraph 5 or 6 further comprising a 5′ UTR, optionally wherein the 5′ UTR comprises the sequence of SEQ ID NO: 2 or SEQ ID NO: 36.

8. The RNA of any one of the preceding paragraphs further comprising a 3′ UTR, optionally wherein the 3′ UTR comprises the sequence of SEQ ID NO: 4 or SEQ ID NO: 37.

9. The RNA of any one of the preceding paragraphs further comprising a 5′ cap analog, optionally a 7mG(5′)ppp(5′)NlmpNp cap.

10. The RNA of any one of the preceding paragraphs further comprising a poly(A) tail, optionally having a length of 50 to 150 nucleotides.

11. The RNA of any one of paragraphs 5-10, wherein the ORF encodes a SARS-CoV-2 antigen.

12. The RNA of paragraph 11, wherein the coronavirus antigen is a structural protein.

13. The RNA of paragraph 12, wherein the structural protein is a spike protein.

14. The RNA of any one of paragraphs 11-13, wherein the coronavirus antigen comprises a sequence having at least 80% identity to the sequence of any one of SEQ ID NOs: 5, 8, 11, 14, 17, 20, 23, 26, 29, 32, 33, 34, 35, 47, 49, 59, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, or 85.

15. The RNA of paragraph 14, wherein the coronavirus antigen comprises a sequence having at least 85%, at least 90%, at least 95%, or at least 98% identity to the sequence of any one of SEQ ID NOs: 5, 8, 11, 14, 17, 20, 23, 26, 29, 32, 33, 34, 35, 47, 49, 59, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, or 85.

16. The RNA of any one of paragraphs 1-13, wherein the ORF comprises the sequence of any one of SEQ ID NOs: 3, 7, 10, 13, 16, 19, 22, 25, 28, 31, 48, 50, 52, 54, 56, 61, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, or 84.

17. The RNA of any one of paragraphs 1-13, wherein the RNA comprises a sequence having at least 85%, at least 90%, at least 95%, or at least 98% identity to the sequence of any one of SEQ ID NOs: 1, 6, 9, 12, 15, 18, 21, 24, 27, 30, 51, 53, 55, 57-58, 60, or 86-97.

18. The RNA of any one of paragraphs 1-13, wherein the RNA comprises the sequence of any one of SEQ ID NOs: 1, 6, 9, 12, 15, 18, 21, 24, 27, 30, 51, 53, 55, 57-58, 60, or 86-97.

19. The RNA of any one of the preceding paragraphs, wherein the RNA comprises a chemical modification and optionally is fully chemically modified.

20. The RNA of paragraph 19, wherein the chemical modification is 1-methylpseudouridine and optionally each uridine is a 1-methylpseudouridine.

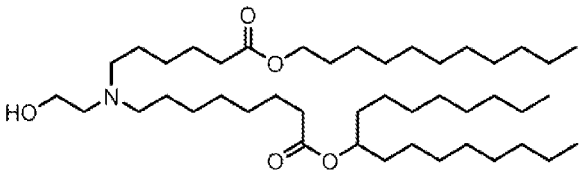
21. The RNA of paragraph 19, wherein each uridine is a 1-methylpseudouridine.

22. The RNA of any one of the preceding paragraphs formulated in a lipid nanoparticle.

23. The RNA of paragraph 22, wherein the lipid nanoparticle comprises a PEG-modified lipid, a non-cationic lipid, a sterol, an ionizable cationic lipid, or any combination thereof.

24. The RNA of paragraph 23, wherein the lipid nanoparticle comprises 0.5-15 mol% PEG-modified lipid; 5-25 mol% non-cationic lipid; 25-55 mol% sterol; and 20-60 mol% ionizable cationic lipid.

25. The RNA of paragraph 24, wherein the PEG-modified lipid is 1,2 dimyristoyl-sn-glycerol, methoxypolyethyleneglycol (PEG2000 DMG), the non-cationic lipid is 1,2 distearoyl-sn-glycero-3-phosphocholine (DSPC), the sterol is cholesterol; and the ionizable cationic lipid has the structure of Compound 1:



(Compound 1).

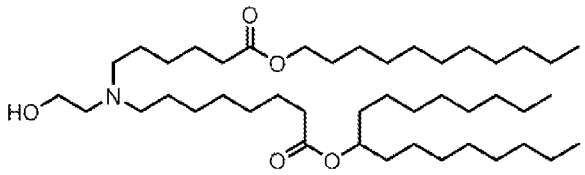
26. A composition comprising the RNA of any one of paragraphs 1-21 and a mixture of lipids.

27. The composition of paragraph 26, wherein the mixture of lipids comprises a PEG-modified lipid, a non-cationic lipid, a sterol, an ionizable cationic lipid, or any combination thereof.

28. The composition of paragraph 27, wherein the mixture of lipids comprises 0.5-15 mol% PEG-modified lipid; 5-25 mol% non-cationic lipid; 25-55 mol% sterol; and 20-60 mol% ionizable cationic lipid.

29. The composition of paragraph 28, wherein the PEG-modified lipid is 1,2 dimyristoyl-sn-glycerol, methoxypolyethyleneglycol (PEG2000 DMG), the non-cationic lipid is 1,2 distearoyl-

sn-glycero-3-phosphocholine (DSPC), the sterol is cholesterol; and the ionizable cationic lipid has the structure of Compound 1:



(Compound 1).

30. The composition of any one of paragraphs 26-29, wherein the mixture of lipids forms lipid nanoparticles.

31. The composition of paragraph 30, wherein the RNA is formulated in the lipid nanoparticles.

32. The RNA of any one of paragraphs 1-13, wherein the ORF comprises a nucleotide sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, or 100% identity to the sequence of SEQ ID NO: 28.

33. The RNA of any one of paragraphs 1-13, wherein the coronavirus antigen comprises an amino acid sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, or 100% identity to the sequence of SEQ ID NO: 29.

33. The RNA of any one of paragraphs 1-13, wherein the RNA comprises a nucleotide sequence having at least 85%, at least 90%, at least 95%, at least 98%, or 100% identity to the sequence of SEQ ID NO: 27.

34. A method comprising administering to a subject the RNA or composition of any one of the preceding paragraphs in an amount effective to induce a neutralizing antibody response against a coronavirus in the subject.

35. A method comprising administering to a subject the RNA or composition of any one of the preceding paragraphs in an amount effective to induce a neutralizing antibody response and/or a T cell immune response, optionally a $CD4^+$ and/or a $CD8^+$ T cell immune response against a coronavirus in the subject.

36. The method of paragraph 34 and 35, wherein the coronavirus is a SARS-CoV-2.

37. The method of any one of the preceding method paragraphs, wherein the subject is immunocompromised.

38. The method of any one of the preceding method paragraphs, wherein the subject has a pulmonary disease.

39. The method of any one of the preceding method paragraphs, wherein the subject is 5 years of age or younger, or 65 years of age or older.

40. The method of any one of the preceding method paragraphs, comprising administering to the subject at least two doses of the composition.

41. The method of any one of the preceding method paragraphs, wherein detectable levels of the coronavirus antigen are produced in serum of the subject at 1-72 hours post administration of the RNA or composition comprising the RNA.

42. The method of any one of the preceding method paragraphs, wherein a neutralizing antibody titer of at least 100 NU/ml, at least 500 NU/ml, or at least 1000 NU/ml is produced in the serum of the subject at 1-72 hours post administration of the RNA or composition comprising the RNA.

43. An immunizing composition comprising: a lipid nanoparticle comprising (a) a messenger RNA that comprises an open reading frame (ORF) having at least 90%, at least 95%, at least 98% or 100% identity to the sequence of SEQ ID NO: 28 and (b) a mixture of lipids comprising 0.5-15 mol% PEG-modified lipid, 5-25 mol% non-cationic lipid, 25-55 mol% sterol, and 20-60 mol% ionizable cationic lipid.

44. An immunizing composition comprising: a lipid nanoparticle comprising (a) a messenger RNA that comprises a sequence having at least 90%, at least 95%, at least 98% or 100% identity to the sequence of SEQ ID NO: 27 and (b) a mixture of lipids comprising 0.5-15 mol% PEG-modified lipid, 5-25 mol% non-cationic lipid, 25-55 mol% sterol, and 20-60 mol% ionizable cationic lipid.

45. An immunizing composition comprising:

(a) a first ribonucleic acid (RNA) 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; and

(b) a second ribonucleic acid (RNA) 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 ORF of the first RNA is different from the ORF of the second RNA.

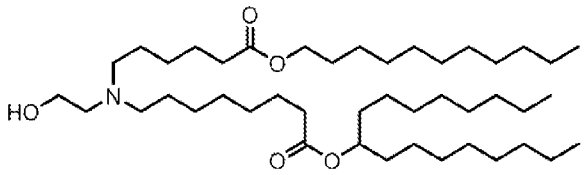
46. The immunizing composition of paragraph 45 further comprising a lipid nanoparticle that comprises a mixture of lipids.

47. The immunizing composition of paragraph 46, wherein the mixture of lipids comprises a PEG-modified lipid, a non-cationic lipid, a sterol, an ionizable cationic lipid, or any combination thereof.

48. The immunizing composition of paragraph 47, wherein the mixture of lipids comprises 0.5-15 mol% PEG-modified lipid; 5-25 mol% non-cationic lipid; 25-55 mol% sterol; and 20-60 mol% ionizable cationic lipid.

49. The immunizing composition of paragraph 46 or 47, wherein the PEG-modified lipid is 1,2 dimyristoyl-sn-glycerol, methoxypolyethyleneglycol (PEG2000 DMG), the non-cationic

lipid is 1,2 distearoyl-sn-glycero-3-phosphocholine (DSPC), the sterol is cholesterol, and the ionizable cationic lipid has the structure of Compound 1:



(Compound 1).

50. The immunizing composition of any one of paragraphs 45-49, wherein coronavirus antigen encoded by the ORF of the first RNA comprises an amino acid sequence having at least 80%, at least 85%, at least 90%, or at least 95% identity to the amino acid sequence of any one of SEQ ID NOs: 5, 8, 11, 14, 17, 20, 23, 26, 29, 32, 33, 34, 35, 47, 49, 59, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, and 85.

51. The immunizing composition of any one of paragraphs 45-49, wherein the ORF of the first RNA comprises a nucleotide sequence having at least 80%, at least 85%, at least 90%, or at least 95% identity to the amino acid sequence of any one of SEQ ID NOs: 3, 7, 10, 13, 16, 19, 22, 25, 28, 31, 48, 50, 52, 54, 56, 61, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, and 84.

52. The RNA of any one of paragraphs 1-13, wherein the ORF encodes a SARS-CoV-2 antigen.

53. The RNA of paragraph 52, wherein the SARS-CoV-2 antigen is a structural protein.

54. The RNA of paragraph 53, wherein the structural protein is selected from the group consisting of: a spike (S) protein, a membrane (M) protein, an envelope (E) protein, and a (NC) nucleocapsid protein.

55. The RNA of paragraph 54, wherein the structural protein is an S protein, optionally a stabilized prefusion form of an S protein.

56. The RNA of paragraph 55, wherein the S protein is an S protein variant, relative to an S protein comprising the amino acid sequence of SEQ ID NO: 32.

57. The RNA of paragraph 56, wherein the S protein variant comprises a reversion of a polybasic cleavage site to a single basic cleavage site.

58. The RNA of paragraph 56, wherein the S protein variant comprises a deletion of a polybasic ER/Golgi signal sequence (KXHXX-COOH) at the carboxy tail of the S protein variant.

59. The RNA of paragraph 57-58, wherein the S protein comprises a double proline stabilizing mutation.

60. The RNA of paragraph 57-58, wherein the S protein comprises a modified protease cleavage site to stabilize the protein.

61. The RNA of paragraph 57-60, wherein the S protein comprises a deletion of the cytoplasmic tail.

62. The RNA of paragraph 57-61, wherein the S protein comprises a foldon scaffold.

63. The RNA of paragraph 57, wherein the S protein comprises a sequence having at least 80% identity to the sequence of any one of SEQ ID NOs: 5, 8, 11, 14, 17, 20, 23, 26, 29, 32, 33, 34, 35, 47, 49, 59, 63, 65, 67, 69, 71, 73, 75, 77, 79, , 81, 83, or 85.

64. The RNA of paragraph 58, wherein the structural protein is an M protein.

65. The RNA of paragraph 64, wherein the M protein comprise a sequence having at least 80% identity to the sequence of SEQ ID NO: 81.

66. The RNA of paragraph 65, wherein the M protein comprises a sequence having at least 85%, at least 90%, at least 95%, or at least 98% identity to the sequence of SEQ ID NO: 81.

67. The RNA of any one of paragraph 57-66, wherein the ORF comprises the sequence of SEQ ID NO: 80.

68. The RNA of any one of paragraph 57-67, wherein the RNA comprises a sequence having at least 85%, at least 90%, at least 95%, or at least 98% identity to the sequence of SEQ ID NO: 95.

69. The RNA of paragraph 68, wherein the RNA comprises the sequence of SEQ ID NO: 95.

70. The RNA of paragraph 54, wherein the structural protein is an E protein.

71. The RNA of paragraph 70, wherein the E protein comprises a sequence having at least 80% identity to the sequence of SEQ ID NO: 83.

72. The RNA of paragraph 71, wherein the E protein comprises a sequence having at least 85%, at least 90%, at least 95%, or at least 98% identity to the sequence of SEQ ID NO: 83.

73. The RNA of any one of paragraph 70-72, wherein the ORF comprises the sequence of SEQ ID NO: 82.

74. The RNA of any one of paragraph 70-73, wherein the RNA comprises a sequence having at least 85%, at least 90%, at least 95%, or at least 98% identity to the sequence of SEQ ID NO: 96.

75. The RNA of paragraph 74, wherein the RNA comprises the sequence of SEQ ID NO: 96.

76. The RNA of paragraph 54, wherein the structural protein is an NC protein.

77. The RNA of paragraph 76, wherein the NC protein comprises a sequence having at least 80% identity to the sequence of SEQ ID NO: 85.

78. The RNA of paragraph 77, wherein the NC protein comprises a sequence having at least 85%, at least 90%, at least 95%, or at least 98% identity to the sequence of SEQ ID NO: 85.

79. The RNA of any one of paragraph 76-78, wherein the ORF comprises the sequence of SEQ ID NO: 84.

80. The RNA of any one of paragraph 76-78, wherein the RNA comprises a sequence having at least 85%, at least 90%, at least 95%, or at least 98% identity to the sequence of SEQ ID NO: 97.

81. The RNA of paragraph 80, wherein the RNA comprises the sequence of SEQ ID NO: 97.

82. The RNA of paragraph 53 wherein the SARS-CoV-2 antigen is a fusion protein.

83. The RNA of paragraph 82, wherein the fusion protein comprises a SARS-CoV-2 polypeptide and a polypeptide from a different virus.

84. A messenger ribonucleic acid (mRNA) comprising an open reading frame (ORF) that comprises a nucleotide sequence having at least 80% identity to the nucleotide sequence of SEQ ID NO 106.

85. A messenger ribonucleic acid (mRNA) comprising a nucleotide sequence having at least 80% identity to the nucleotide sequence of SEQ ID NO 105.

86. A messenger ribonucleic acid (mRNA) comprising an open reading frame (ORF) that comprises a nucleotide sequence having at least 95% identity to the nucleotide sequence of SEQ ID NO 106.

87 A messenger ribonucleic acid (mRNA) comprising a nucleotide sequence having at least 95% identity to the nucleotide sequence of SEQ ID NO 105.

88. A messenger ribonucleic acid (mRNA) comprising an open reading frame (ORF) that comprises a nucleotide sequence having at least 99% identity to the nucleotide sequence of SEQ ID NO 106.

89. A messenger ribonucleic acid (mRNA) comprising a nucleotide sequence having at least 99% identity to the nucleotide sequence of SEQ ID NO 105.

SEQUENCE LISTING

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: 36)

5′ UTR: GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCGGCGCCGCCACC (SEQ ID NO: 2)

3′ UTR: UGAUAAUAGGCUGGAGCCUCGGUGGCCAUGCUUCUUGCCCCUUGGGCCUCCCCCCAGCC CCUCCUCCCCUUCCUGCACCCGUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC (SEQ ID NO: 37)

3′ UTR: UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGGCCUCCCCCCAGCC CCUCCUCCCCUUCCUGCACCCGUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC (SEQ ID NO: 4)

Table 1.

SARS-CoV-2 **Spike (S) Protein**		
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	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUA AGACCCCGGCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGUGAGCAGC CAGUGCGUGAACCUGACCACCCGGACCCAGCUGCCACCAG CCUACACCAACAGCUUCACCCGGGGCGUCUACUACCCCGA CAAGGUGUUCCGGAGCAGCGUCCUGCACAGCACCCAGGA CCUGUUCCUGCCCUUCUUCAGCAACGUGACCUGGUUCCAC GCCAUCCACGUGAGCGGCACCAACGGCACCAAGCGGUUCG ACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUCGC CAGCACCGAGAAGAGCAACAUCAUCCGGGGCUGGAUCUU CGGCACCACCCUGGACAGCAAGACCCAGAGCCUGCUGAUC GUGAAUAACGCCACCAACGUGGUGAUCAAGGUGUGCGAG UUCCAGUUCUGCAACGACCCCUUCCUGGGCGUGUACUACC ACAAGAACAACAAGAGCUGGAUGGAGAGCGAGUUCCGGG UGUACAGCAGCGCCAACAACUGCACCUUCGAGUACGUGA GCCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAGGGCA ACUUCAAGAACCUGCGGGAGUUCGUGUUCAAGAACAUCG ACGGCUACUUCAAGAUCUACAGCAAGCACACCCCAAUCA ACCUGGUGCGGGAUCUGCCCCAGGGCUUCUCAGCCCUGG AGCCCCUGGUGGACCUGCCCAUCGGCAUCAACAUCACCCG GUUCCAGACCCUGCUGGCCCUGCACCGGAGCUACCUGACC CCAGGCGACAGCAGCAGCGGGUGGACAGCAGGCGCGGCU GCUUACUACGUGGGCUACCUGCAGCCCCGGACCUUCCUGC UGAAGUACAACGAGAACGGCACCAUCACCGACGCCGUGG ACUGCGCCCUGGACCCUCUGAGCGAGACCAAGUGCACCCU GAAGAGCUUCACCGUGGAGAAGGGCAUCUACCAGACCAG CAACUUCCGGGUGCAGCCCACCGAGAGCAUCGUGCGGUU CCCCAACAUCACCAACCUGUGCCCCUUCGGCGAGGUGUUC AACGCCACCCGGUUCGCCAGCGUGUACGCCUGGAACCGGA AGCGGAUCAGCAACUGCGUGGCCGACUACAGCGUGCUGU ACAACAGCGCCAGCUUCAGCACCUUCAAGUGCUACGGCG UGAGCCCCACCAAGCUGAACGACCUGUGCUUCACCAACGU GUACGCCGACAGCUUCGUGAUCCGUGGCGACGAGGUGCG	31

	GCAGAUCGCACCCGGCCAGACAGGCAAGAUCGCCGACUAC AACUACAAGCUGCCCGACGACUUCACCGGCUGCGUGAUC GCCUGGAACAGCAACAACCUCGACAGCAAGGUGGGCGGC AACUACAACUACCUGUACCGGCUGUUCCGGAAGAGCAAC CUGAAGCCCUUCGAGCGGGACAUCAGCACCGAGAUCUAC CAAGCCGGCUCCACCCCUUGCAACGGCGUGGAGGGCUUCA ACUGCUACUUCCCUCUGCAGAGCUACGGCUUCCAGCCCAC CAACGGCGUGGGCUACCAGCCCUACCGGGUGGUGGUGCU GAGCUUCGAGCUGCUGCACGCCCCAGCCACCGUGUGUGGC CCCAAGAAGAGCACCAACCUGGUGAAGAACAAGUGCGUG AACUUCAACUUCAACGGCCUUACCGGCACCGGCGUGCUG ACCGAGAGCAACAAGAAAUUCCUGCCCUUUCAGCAGUUC GGCCGGGACAUCGCCGACACCACCGACGCUGUGCGGGAUC CCCAGACCCUGGAGAUCCUGGACAUCACCCCUUGCAGCUU CGGCGGCGUGAGCGUGAUCACCCCAGGCACCAACACCAGC AACCAGGUGGCCGUGCUGUACCAGGACGUGAACUGCACC GAGGUGCCCGUGGCCAUCCACGCCGACCAGCUGACACCCA CCUGGCGGGUCUACAGCACCGGCAGCAACGUGUUCCAGA CCCGGGCCGGUUGCCUGAUCGGCGCCGAGCACGUGAACA ACAGCUACGAGUGCGACAUCCCCAUCGGCGCCGGCAUCUG UGCCAGCUACCAGACCCAGACCAAUUCACCCCGGAGGGCA AGGAGCGUGGCCAGCCAGAGCAUCAUCGCCUACACCAUG AGCCUGGGCGCCGAGAACAGCGUGGCCUACAGCAACAAC AGCAUCGCCAUCCCCACCAACUUCACCAUCAGCGUGACCA CCGAGAUUCUGCCCGUGAGCAUGACCAAGACCAGCGUGG ACUGCACCAUGUACAUCUGCGGCGACAGCACCGAGUGCA GCAACCUGCUGCUGCAGUACGGCAGCUUCUGCACCCAGCU GAACCGGGCCCUGACCGGCAUCGCCGUGGAGCAGGACAA GAACACCCAGGAGGUGUUCGCCCAGGUGAAGCAGAUCUA CAAGACCCCUCCCAUCAAGGACUUCGGCGGCUUCAACUUC AGCCAGAUCCUGCCCGACCCCAGCAAGCCCAGCAAGCGGA GCUUCAUCGAGGACCUGCUGUUCAACAAGGUGACCCUAG CCGACGCCGGCUUCAUCAAGCAGUACGGCGACUGCCUCGG CGACAUAGCCGCCCGGGACCUGAUCUGCGCCCAGAAGUUC AACGGCCUGACCGUGCUGCCUCCCCUGCUGACCGACGAGA UGAUCGCCCAGUACACCAGCGCCCUGUUAGCCGGAACCAU CACCAGCGGCUGGACUUUCGGCGCUGGAGCCGCUCUGCA GAUCCCCUUCGCCAUGCAGAUGGCCUACCGGUUCAACGGC AUCGGCGUGACCCAGAACGUGCUGUACGAGAACCAGAAG CUGAUCGCCAACCAGUUCAACAGCGCCAUCGGCAAGAUCC AGGACAGCCUGAGCAGCACCGCUAGCGCCCUGGGCAAGC UGCAGGACGUGGUGAACCAGAACGCCCAGGCCCUGAACA CCCUGGUGAAGCAGCUGAGCAGCAACUUCGGCGCCAUCA GCAGCGUGCUGAACGACAUCCUGAGCCGGCUGGACAAGG UGGAGGCCGAGGUGCAGAUCGACCGGCUGAUCACUGGCC GGCUGCAGAGCCUGCAGACCUACGUGACCCAGCAGCUGA UCCGGGCCGCCGAGAUUCGGGCCAGCGCCAACCUGGCCGC CACCAAGAUGAGCGAGUGCGUGCUGGGCCAGAGCAAGCG GGUGGACUUCUGCGGCAAGGGCUACCACCUGAUGAGCUU UCCCCAGAGCGCACCCCACGGAGUGGUGUUCCUGCACGUG ACCUACGUGCCCGCCCAGGAGAAGAACUUCACCACCGCCC CAGCCAUCUGCCACGACGGCAAGGCCCACUUUCCCCGGGA GGGCGUGUUCGUGAGCAACGGCACCCACUGGUUCGUGAC CCAGCGGAACUUCUACGAGCCCCAGAUCAUCACCACCGAC AACACCUUCGUGAGCGGCAACUGCGACGUGGUGAUCGGC AUCGUGAACAACACCGUGUACGAUCCCCUGCAGCCCGAGC UGGACAGCUUCAAGGAGGAGCUGGACAAGUACUUCAAGA AUCACACCAGCCCCGACGUGGACCUGGGCGACAUCAGCGG	

	CAUCAACGCCAGCGUGGUGAACAUCCAGAAGGAGAUCGA UCGGCUGAACGAGGUGGCCAAGAACCUGAACGAGAGCCU GAUCGACCUGCAGGAGCUGGGCAAGUACGAGCAGUACAU CAAGUGGCCCUGGUACAUCUGGCUGGGCUUCAUCGCCGG CCUGAUCGCCAUCGUGAUGGUGACCAUCAUGCUGUGCUG CAUGACCAGCUGCUGCAGCUGCCUGAAGGGCUGUUGCAG CUGCGGCAGCUGCUGCAAGUUCGACGAGGACGACAGCGA GCCCGUGCUGAAGGGCGUGAAGCUGCACUACACC	
3′ UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCC CCUUGGGCCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACC CGUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGG C	4
Corresponding amino acid sequence	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVF RSSVLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPF NDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKV CEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVS QPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRD LPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWT AGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKC TLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATR FASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLN DLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTG CVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQ AGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFE LLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKK FLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTN TSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQT RAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPRRARSVA SQSIIAYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTK TSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDK NTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDL LFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPL LTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRF NGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQ DVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEV QIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVL GQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNF TTAPAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTD NTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTS PDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGK YEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCLKGCC SCGSCCKFDEDDSEPVLKGVKLHYT	32
PolyA tail	100 nt	
SARS-CoV-2 S protein Variant 1		
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′)NlmpNp	
5′ UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUA AGACCCCGGCGCCGCCACC	2
ORF of mRNA	AUGUUCGUCUUCCUCGUCUUGCUGCCGCUGGUGUCGAGC CAGUGCGUGAACCUCACCACAAGGACGCAGCUCCCACCGG	3

Construct (excluding the stop codon)	CCUACACGAACAGCUUCACGCGCGGCGUGUACUACCCCGA CAAGGUGUUCCGGUCGUCGGUCCUCCACUCCACGCAGGAC CUCUUCCUGCCCUUCUUCAGCAACGUGACCUGGUUCCACG CCAUCCACGUCUCCGGGACGAACGGGACGAAGCGGUUCG ACAACCCGGUCCUCCCGUUCAACGACGGCGUCUACUUCGC GAGCACGGAGAAGUCGAACAUCAUCCGGGGCUGGAUCUU CGGCACGACCCUGGACUCGAAGACCCAGUCCCUACUUAUC GUGAACAACGCCACCAACGUCGUCAUCAAGGUCUGCGAG UUCCAGUUCUGCAACGACCCCUUCCUCGGCGUCUACUACC ACAAGAACAACAAGUCGUGGAUGGAGUCGGAGUUCCGGG UGUACAGCUCGGCGAACAACUGCACCUUCGAGUACGUGU CGCAGCCGUUCCUCAUGGACCUCGAGGGCAAGCAGGGUA ACUUCAAGAACCUGCGCGAGUUCGUCUUCAAGAACAUCG ACGGCUACUUCAAGAUCUACUCCAAGCACACGCCCAUCAA CCUGGUCCGCGACCUCCCGCAAGGCUUCUCCGCCCUCGAG CCUCUGGUCGACCUGCCGAUCGGCAUCAACAUCACGAGG UUCCAGACGCUCCUGGCGCUGCACCGGUCGUACCUGACGC CAGGCGACUCCUCCUCGGGCUGGACAGCAGGCGCGGCUGC CUACUACGUCGGGUACCUGCAGCCCCGCACGUUCCUCCUG AAGUACAACGAGAACGGCACUAUCACGGACGCCGUCGAC UGCGCCCUGGACCCACUGUCGGAGACGAAGUGCACGCUG AAGUCGUUCACCGUGGAGAAGGGUAUCUACCAGACCUCC AACUUCCGGGUCCAGCCGACGGAGUCGAUCGUGCGGUUC CCCAACAUCACGAACCUGUGCCCCUUCGGUGAGGUCUUCA ACGCCACCCGGUUCGCGUCGGUCUACGCGUGGAACCGUA AGCGCAUCUCGAACUGCGUGGCGGACUACUCCGUCCUCU ACAACAGCGCGUCCUUCAGCACCUUCAAGUGCUACGGCG UCAGCCCCACGAAGCUGAACGACCUCUGCUUCACCAACGU CUACGCAGACUCCUUCGUGAUCCGGGGUGACGAGGUGCG ACAGAUCGCCCCUGGUCAGACCGGGAAGAUCGCCGACUA CAACUACAAGCUGCCCGACGACUUCACCGGCUGCGUGAUC GCGUGGAACAGCAACAACCUGGACUCCAAGGUCGGAGGU AACUACAACUACCUCUACCGGCUGUUCCGCAAGUCCAACC UGAAGCCGUUCGAGCGGGACAUCUCCACGGAGAUCUACC AAGCCGGCUCGACCCCUUGUAACGGGGUGGAGGGGUUCA ACUGCUACUUCCCACUGCAGUCCUACGGGUUCCAGCCCAC CAACGGGGUCGGGUACCAGCCGUACCGCGUGGUGGUCCU GUCCUUCGAGCUGCUGCACGCGCCAGCCACGGUGUGCGG GCCAAAGAAGAGCACGAACCUGGUCAAGAACAAGUGCGU CAACUUCAACUUCAACGGCCUGACGGGGACAGGGGUCCU CACGGAGUCGAACAAGAAGUUCCUGCCGUUCCAGCAGUU CGGCCGUGACAUCGCAGACACGACUGACGCCGUCCGCGAC CCUCAGACCCUCGAGAUCCUCGACAUCACCCCGUGCUCGU UCGGCGGAGUGAGCGUCAUCACCCCGGGGACCAACACAU CGAACCAGGUGGCCGUCCUGUACCAGGACGUCAACUGCA CGGAGGUCCCUGUGGCGAUCCACGCCGACCAGCUCACGCC CACCUGGCGCGUCUACUCCACCGGGUCCAACGUGUUCCAG ACCCGCGCAGGCUGCCUGAUCGGGGCCGAGCACGUCAACA ACAGCUACGAGUGCGACAUCCCCAUCGGAGCGGGCAUCU GCGCCAGCUACCAGACGCAGACGAACUCUCCAAGGCGCGC UCGUAGCGUGGCCUCCCAGUCCAUCAUCGCGUACACGAU GUCCCUUGGGGCCGAGAACUCGGUCGCAUACAGCAACAA CUCCAUCGCCAUCCCCACCAACUUCACGAUCUCGGUCACC ACCGAGAUCCUCCCGGUCAGCAUGACGAAGACGUCGGUG GACUGCACCAUGUACAUCUGCGGGGACAGCACGGAGUGC UCGAACCUGCUCCUGCAGUACGGGAGCUUCUGCACCCAGC UGAACAGGGCGCUGACGGGGAUCGCGGUGGAGCAGGACA AGAACACCCAGGAGGUGUUCGCGCAGGUGAAGCAGAUCU	

	ACAAGACGCCUCCAAUCAAGGACUUCGGCGGGUUCAACU UCUCGCAGAUCCUCCCCGACCCGUCCAAGCCGUCGAAGCG GUCGUUCAUCGAGGACCUGCUCUUCAACAAGGUGACGUU GGCCGACGCGGGCUUCAUCAAGCAGUACGGGGACUGCCU UGGGGACAUCGCUGCCCGCGACCUCAUCUGCGCCCAGAAG UUCAACGGGCUGACUGUGCUCCCGCCCCUGCUGACGGACG AGAUGAUCGCGCAGUACACGUCCGCGCUGCUCGCUGGAA CGAUCACCUCCGGGUGGACCUUCGGCGCUGGAGCGGCUC UGCAGAUCCCGUUCGCGAUGCAGAUGGCGUACCGGUUCA ACGGCAUCGGGGUGACCCAGAACGUCCUCUACGAGAACC AGAAGCUGAUCGCCAACCAGUUCAACUCCGCGAUCGGCA AGAUCCAGGACUCGCUGAGCUCCACGGCUUCCGCCCUCGG GAAGCUUCAGGACGUGGUGAACCAGAACGCCCAGGCCCU CAACACCCUGGUGAAGCAGCUGAGCUCGAACUUCGGCGC CAUCUCGAGCGUGCUCAACGACAUCCUGAGCCGUCUGGA CCCUCCCGAGGCGGAGGUGCAGAUCGACCGGCUCAUCACG GGCCGGCUUCAGUCCCUGCAGACGUACGUGACCCAGCAGC UCAUACGGGCGGCGGAGAUACGCGCCUCCGCCAACCUGGC CGCGACGAAGAUGUCCGAGUGCGUCCUCGGACAGAGCAA GCGCGUGGACUUCUGCGGCAAGGGGUACCACCUCAUGAG CUUUCCCCAGUCGGCUCCUCACGGGGUCGUCUUCCUGCAC GUGACGUACGUCCCGGCGCAGGAGAAGAACUUCACCACC GCCCCAGCGAUCUGCCACGACGGGAAGGCGCACUUCCCGC GCGAGGGCGUCUUCGUCUCCAACGGGACCCACUGGUUCG UCACCCAGCGGAACUUCUACGAGCCGCAGAUCAUCACGAC CGACAACACGUUCGUAUCCGGGAACUGCGACGUCGUCAU CGGCAUCGUCAACAACACGGUCUACGACCCACUGCAGCCG GAGCUGGACUCGUUCAAGGAGGAGCUGGACAAGUAUUUC AAGAACCACACCUCGCCCGACGUGGACCUGGGCGACAUCA GCGGGAUCAACGCGUCGGUCGUGAACAUCCAGAAGGAGA UCGACCGACUGAACGAGGUCGCCAAGAACCUGAACGAGU CCCUGAUCGACCUGCAAGAGCUCGGCAAGUACGAGCAGU ACAUCAAGUGGCCUUGGUACAUCUGGCUCGGCUUCAUCG CGGGGCUGAUCGCCAUCGUGAUGGUCACCAUCAUGUUGU GCUGCAUGACCUCCUGCUGCUCGUGCCUCAAGGGGUGCU GCAGCUGCGGGUCCUGCUGCAAGUUCGACGAGGACGACU CGGAGCCGGUCCUCAAGGGCGUCAAGCUCCACUACACC	
3′ UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCC CCUUGGGCCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACC CGUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGG C	4
Corresponding amino acid sequence	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVF RSSVLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPF NDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKV CEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVS QPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRD LPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWT AGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKC TLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATR FASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLN DLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTG CVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQ AGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFE LLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKK FLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTN TSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQT RAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPRRARSVA	5

	SQSIIAYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTK TSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDK NTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDL LFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPL LTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRF NGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQ DVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDPPEAEVQ IDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLG QSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTT APAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNT FVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPD VDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYE QYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCLKGCCSC GSCCKFDEDDSEPVLKGVKLHYT	
PolyA tail	100 nt	
SARS-CoV-2 S Protein Variant 2		
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	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUA AGACCCCGGCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGUGAGCAGC CAGUGCGUGAACCUGACCACCAGGACCCAGCUGCCGCCUG CCUACACCAACAGCUUCACCCGCGGUGUGUACUACCCCGA CAAGGUGUUCAGGUCCAGCGUGCUGCACAGCACCCAGGA CCUGUUCCUCCCCUUCUUCAGCAACGUGACCUGGUUCCAC GCCAUCCACGUGAGCGGCACCAACGGCACCAAGCGGUUCG ACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUCGC CAGCACCGAGAAGAGCAACAUCAUCCGGGGCUGGAUCUU CGGCACCACACUCGACAGCAAGACCCAGAGCCUGCUGAUC GUGAACAACGCCACCAACGUGGUGAUCAAGGUGUGCGAA UUCCAGUUCUGCAACGACCCCUUCCUGGGCGUGUACUACC ACAAGAACAACAAGAGCUGGAUGGAGAGCGAGUUCCGCG UGUACAGCAGCGCCAACAACUGCACCUUCGAGUACGUGA GCCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAGGGCA AUUUCAAGAACCUGAGGGAGUUCGUGUUCAAGAACAUCG ACGGCUACUUCAAGAUCUACAGCAAGCACACGCCCAUCA ACCUGGUGCGGGACUUGCCCCAGGGCUUCAGCGCCCUGG AGCCCUUAGUGGACCUGCCUAUCGGCAUCAACAUCACCCG GUUCCAGACCCUGCUGGCCCUGCACCGGAGCUACCUGACU CCCGGCGACAGCAGCUCCGGGUGGACUGCCGGUGCUGCCG CCUACUACGUGGGGUACCUGCAGCCCCGGACCUUCCUGCU GAAGUACAACGAGAACGGCACCAUCACCGACGCCGUGGA CUGCGCCCUGGAUCCACUGAGCGAGACCAAGUGCACCCUG AAGAGCUUCACCGUGGAGAAGGGCAUCUACCAGACCAGC AACUUCCGGGUGCAGCCCACCGAGAGCAUCGUGAGGUUC CCCAACAUCACCAACCUGUGCCCUUUCGGCGAGGUGUUCA ACGCCACCCGCUUCGCCUCCGUGUACGCCUGGAACAGGAA GAGGAUCAGCAACUGCGUGGCCGACUACAGCGUGCUGUA CAACAGCGCCAGCUUCUCCACCUUCAAGUGCUACGGCGUG AGCCCAACCAAGCUGAACGACCUGUGCUUUACCAACGUG UACGCCGAUAGCUUCGUGAUCCGCGGCGACGAAGUGCGG	7

	CAGAUCGCUCCUGGGCAGACCGGAAAGAUCGCCGACUAC AACUACAAGCUGCCCGACGACUUCACCGGGUGCGUGAUC GCUUGGAACAGCAACAACCUGGACAGCAAGGUGGGCGGC AACUACAACUACCUGUACCGGCUGUUCCGGAAGAGCAAC CUGAAGCCCUUCGAGCGCGACAUCUCCACCGAGAUCUACC AGGCCGGCUCCACACCCUGCAACGGCGUGGAGGGCUUCA ACUGCUACUUUCCCCUGCAGUCCUACGGCUUCCAGCCCAC CAACGGCGUGGGCUACCAGCCAUACCGCGUGGUGGUGCU GUCCUUCGAGCUGCUGCACGCUCCCGCCACCGUUUGCGGC CCCAAGAAGUCCACCAACCUGGUGAAGAACAAGUGCGUG AACUUCAACUUCAACGGUCUCACGGGCACCGGGGUGCUG ACCGAGAGCAACAAGAAGUUCCUGCCCUUUCAGCAGUUC GGCAGGGACAUCGCCGACACCACAGACGCCGUGCGGGAU CCCCAGACCCUGGAGAUCCUGGACAUCACCCCGUGCAGCU UCGGCGGCGUGAGCGUGAUCACGCCCGGCACCAACACCAG CAACCAGGUGGCCGUGCUGUACCAGGACGUGAACUGCAC CGAGGUGCCCGUGGCCAUCCACGCCGACCAGCUGACUCCC ACCUGGCGCGUGUAUAGCACCGGCAGCAACGUGUUCCAG ACACGGGCCGGCUGCCUGAUCGGCGCCGAGCACGUGAAC AACUCCUACGAGUGCGACAUCCCCAUCGGCGCUGGCAUCU GCGCCAGCUACCAGACCCAGACCAACAGCCCCAGACGGGC CAGGUCCGUGGCUUCCCAGAGCAUCAUCGCCUACACCAUG UCCCUGGGCGCCGAGAACAGCGUGGCCUACAGCAACAAC UCCAUCGCCAUCCCCCACCAACUUCACCAUCAGCGUGACCA CCGAGAUCCUGCCCGUGAGCAUGACCAAGACCUCCGUGG ACUGCACCAUGUACAUCUGCGGCGACAGCACCGAGUGCA GCAACCUGCUGCUGCAGUACGGCAGCUUCUGCACCCAGCU GAACAGGGCCCUGACCGGCAUCGCCGUGGAGCAGGACAA GAACACCCAGGAGGUGUUCGCCCAGGUGAAGCAGAUCUA CAAGACUCCACCUAUCAAGGACUUCGGCGGGUUCAACUU CAGCCAGAUCCUCCCCGACCCCUCCAAGCCCAGCAAGCGG AGCUUCAUCGAGGACCUGCUGUUCAACAAGGUGACCCUG GCUGACGCCGGCUUUAUCAAGCAGUACGGCGACUGCCUU GGCGACAUCGCCGCCAGGGACCUGAUCUGCGCCCAGAAG UUCAACGGCCUGACCGUGCUGCCGCCACUGCUGACCGACG AGAUGAUCGCCCAGUACACCUCUGCCCUGCUGGCCGGUAC CAUCACCUCCGGCUGGACAUUUGGUGCUGGCGCUGCGCU GCAGAUCCCCUUCGCCAUGCAGAUGGCCUACCGCUUCAAC GGCAUCGGGGUGACCCAGAACGUGCUGUACGAGAACCAG AAGCUGAUCGCCAACCAGUUCAACAGCGCCAUCGGCAAG AUCCAGGACAGCCUGAGCAGCACCGCCAGCGCUCUGGGCA AGCUGCAGGACGUGGUGAACCAGAACGCCCAGGCCCUGA ACACCCUGGUGAAGCAGCUGUCCAGCAACUUCGGCGCCA UCAGCUCCGUGCUGAACGACAUCCUGAGCCGGCUGGAUC CACCAGAGGCCGAGGUGCAGAUCGACCGUCUGAUCACCG GUCGGCUGCAGAGCCUGCAGACCUACGUGACCCAGCAGC UGAUCCGCGCCGCCGAAAUCCGCGCCUCCGCCAACCUGGC CGCCACCAAGAUGUCCGAGUGCGUGCUGGGCCAGAGCAA GCGGGUGGACUUCUGCGGCAAGGGCUACCACCUGAUGAG CUUCCCACAGAGCGCUCCCCACGGGGUAGUGUUCCUGCAC GUGACCUACGUGCCCGCCCAGGAGAAGAACUUCACCACU GCACCCGCCAUCUGCCACGACGGCAAGGCCCACUUCCCUC GGGAGGGCGUGUUCGUGAGCAACGGCACCCACUGGUUCG UGACCCAGAGGAACUUCUACGAGCCCCAGAUCAUCACCAC CGACAACACCUUCGUGUCCGGCAACUGCGACGUGGUGAU CGGCAUAGUGAACAACACCGUGUACGACCCACUGCAGCCC GAGCUGGACAGCUUCAAGGAGGAGCUGGACAAGUACUUC AAGAACCACACCAGCCCAGACGUGGACCUGGGCGACAUC	

	UCCGGCAUCAACGCCUCCGUGGUGAACAUCCAGAAGGAG AUCGACCGGCUGAACGAGGUGGCCAAGAACCUGAACGAG AGCCUGAUCGACCUGCAGGAGCUGGGGAAGUACGAGCAG UACAUCAAGUGGCCUUGGUACAUCUGGCUGGGCUUCAUC GCCGGCCUGAUCGCCAUCGUGAUGGUGACCAUCAUGCUG UGCUGCAUGACCAGCUGCUGCAGCUGCCUGAAGGGCUGU UGCAGCUGCGGCAGCUGCUGCAAGUUCGACGAGGACGAC AGCGAGCCCGUGCUGAAGGGCGUGAAGCUGCACUACACC	
3′ UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCC CCUUGGGCCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACC CGUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGG C	4
Corresponding amino acid sequence	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVF RSSVLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPF NDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKV CEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVS QPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRD LPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWT AGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKC TLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATR FASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLN DLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTG CVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQ AGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFE LLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKK FLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTN TSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQT RAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPRRARSVA SQSIIAYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTK TSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDK NTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDL LFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPL LTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRF NGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQ DVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDPPEAEVQ IDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLG QSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTT APAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNT FVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPD VDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYE QYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCLKGCCSC GSCCKFDEDDSEPVLKGVKLHYT	8
PolyA tail	100 nt	
SARS-CoV-2 S Protein Variant 3		
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	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUA AGACCCCGGCGCCGCCACC	2
ORF of mRNA	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGUGAGCAGC CAGUGCGUGAACCUGACCACCCGGACCCAGCUGCCACCAG	10

Construct (excluding the stop codon)	CCUACACCAACAGCUUCACCCGGGGCGUCUACUACCCCGA CAAGGUGUUCCGGAGCAGCGUCCUGCACAGCACCCAGGA CCUGUUCCUGCCCUUCUUCAGCAACGUGACCUGGUUCCAC GCCAUCCACGUGAGCGGCACCAACGGCACCAAGCGGUUCG ACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUCGC CAGCACCGAGAAGAGCAACAUCAUCCGGGGCUGGAUCUU CGGCACCACCCUGGACAGCAAGACCCAGAGCCUGCUGAUC GUGAAUAACGCCACCAACGUGGUGAUCAAGGUGUGCGAG UUCCAGUUCUGCAACGACCCCUUCCUGGGCGUGUACUACC ACAAGAACAACAAGAGCUGGAUGGAGAGCGAGUUCCGGG UGUACAGCAGCGCCAACAACUGCACCUUCGAGUACGUGA GCCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAGGGCA ACUUCAAGAACCUGCGGGAGUUCGUGUUCAAGAACAUCG ACGGCUACUUCAAGAUCUACAGCAAGCACACCCCAAUCA ACCUGGUGCGGGAUCUGCCCCAGGGCUUCUCAGCCCUGG AGCCCCUGGUGGACCUGCCCAUCGGCAUCAACAUCACCCG GUUCCAGACCCUGCUGGCCCUGCACCGGAGCUACCUGACC CCAGGCGACAGCAGCAGCGGGUGGACAGCAGGCGCGGCU GCUUACUACGUGGGCUACCUGCAGCCCCGGACCUUCCUGC UGAAGUACAACGAGAACGGCACCAUCACCGACGCCGUGG ACUGCGCCCUGGACCCUCUGAGCGAGACCAAGUGCACCCU GAAGAGCUUCACCGUGGAGAAGGGCAUCUACCAGACCAG CAACUUCCGGGUGCAGCCCACCGAGAGCAUCGUGCGGUU CCCCAACAUCACCAACCUGUGCCCCUUCGGCGAGGUGUUC AACGCCACCCGGUUCGCCAGCGUGUACGCCUGGAACCGGA AGCGGAUCAGCAACUGCGUGGCCGACUACAGCGUGCUGU ACAACAGCGCCAGCUUCAGCACCUUCAAGUGCUACGGCG UGAGCCCCACCAAGCUGAACGACCUGUGCUUCACCAACGU GUACGCCGACAGCUUCGUGAUCCGUGGCGACGAGGUGCG GCAGAUCGCACCCGGCCAGACAGGCAAGAUCGCCGACUAC AACUACAAGCUGCCCGACGACUUCACCGGCUGCGUGAUC GCCUGGAACAGCAACAACCUCGACAGCAAGGUGGGCGGC AACUACAACUACCUGUACCGGCUGUUCCGGAAGAGCAAC CUGAAGCCCUUCGAGCGGGACAUCAGCACCGAGAUCUAC CAAGCCGGCUCCACCCCUUGCAACGGCGUGGAGGGCUUCA ACUGCUACUUCCCUCUGCAGAGCUACGGCUUCCAGCCCAC CAACGGCGUGGGCUACCAGCCCUACCGGGUGGUGGUGCU GAGCUUCGAGCUGCUGCACGCCCCAGCCACCGUGUGUGGC CCCAAGAAGAGCACCAACCUGGUGAAGAACAAGUGCGUG AACUUCAACUUCAACGGCCUUACCGGCACCGGCGUGCUG ACCGAGAGCAACAAGAAAUUCCUGCCCUUUCAGCAGUUC GGCCGGGACAUCGCCGACACCACCGACGCUGUGCGGGAUC CCCAGACCCUGGAGAUCCUGGACAUCACCCCUUGCAGCUU CGGCGGCGUGAGCGUGAUCACCCCAGGCACCAACACCAGC AACCAGGUGGCCGUGCUGUACCAGGACGUGAACUGCACC GAGGUGCCCGUGGCCAUCCACGCCGACCAGCUGACACCCA CCUGGCGGGUCUACAGCACCGGCAGCAACGUGUUCCAGA CCCGGGCCGGUUGCCUGAUCGGCGCCGAGCACGUGAACA ACAGCUACGAGUGCGACAUCCCCAUCGGCGCCGGCAUCUG UGCCAGCUACCAGACCCAGACCAAUUCACCCGGCAGCGGC GGCAGCGUGGCCAGCCAGAGCAUCAUCGCCUACACCAUG AGCCUGGGCGCCGAGAACAGCGUGGCCUACAGCAACAAC AGCAUCGCCAUCCCCACCAACUUCACCAUCAGCGUGACCA CCGAGAUUCUGCCCGUGAGCAUGACCAAGACCAGCGUGG ACUGCACCAUGUACAUCUGCGGCGACAGCACCGAGUGCA GCAACCUGCUGCUGCAGUACGGCAGCUUCUGCACCCAGCU GAACCGGGCCCUGACCGGCAUCGCCGUGGAGCAGGACAA GAACACCCAGGAGGUGUUCGCCCAGGUGAAGCAGAUCUA	

	CAAGACCCCUCCCAUCAAGGACUUCGGCGGCUUCAACUUC AGCCAGAUCCUGCCCGACCCCAGCAAGCCCAGCAAGCGGA GCUUCAUCGAGGACCUGCUGUUCAACAAGGUGACCCUAG CCGACGCCGGCUUCAUCAAGCAGUACGGCGACUGCCUCGG CGACAUAGCCGCCCGGGACCUGAUCUGCGCCCAGAAGUUC AACGGCCUGACCGUGCUGCCUCCCCUGCUGACCGACGAGA UGAUCGCCCAGUACACCAGCGCCCUGUUAGCCGGAACCAU CACCAGCGGCUGGACUUUCGGCGCUGGAGCCGCUCUGCA GAUCCCCUUCGCCAUGCAGAUGGCCUACCGGUUCAACGGC AUCGGCGUGACCCAGAACGUGCUGUACGAGAACCAGAAG CUGAUCGCCAACCAGUUCAACAGCGCCAUCGGCAAGAUCC AGGACAGCCUGAGCAGCACCGCUAGCGCCCUGGGCAAGC UGCAGGACGUGGUGAACCAGAACGCCCAGGCCCUGAACA CCCUGGUGAAGCAGCUGAGCAGCAACUUCGGCGCCAUCA GCAGCGUGCUGAACGACAUCCUGAGCCGGCUGGACCCUCC CGAGGCCGAGGUGCAGAUCGACCGGCUGAUCACUGGCCG GCUGCAGAGCCUGCAGACCUACGUGACCCAGCAGCUGAU CCGGGCCGCCGAGAUUCGGGCCAGCGCCAACCUGGCCGCC ACCAAGAUGAGCGAGUGCGUGCUGGGCCAGAGCAAGCGG GUGGACUUCUGCGGCAAGGGCUACCACCUGAUGAGCUUU CCCCAGAGCGCACCCCACGGAGUGGUGUUCCUGCACGUGA CCUACGUGCCCGCCCAGGAGAAGAACUUCACCACCGCCCC AGCCAUCUGCCACGACGGCAAGGCCCACUUUCCCCGGGAG GGCGUGUUCGUGAGCAACGGCACCCACUGGUUCGUGACC CAGCGGAACUUCUACGAGCCCCAGAUCAUCACCACCGACA ACACCUUCGUGAGCGGCAACUGCGACGUGGUGAUCGGCA UCGUGAACAACACCGUGUACGAUCCCCUGCAGCCCGAGCU GGACAGCUUCAAGGAGGAGCUGGACAAGUACUUCAAGAA UCACACCAGCCCCGACGUGGACCUGGGCGACAUCAGCGGC AUCAACGCCAGCGUGGUGAACAUCCAGAAGGAGAUCGAU CGGCUGAACGAGGUGGCCAAGAACCUGAACGAGAGCCUG AUCGACCUGCAGGAGCUGGGCAAGUACGAGCAGGGCAGC GGCUACAUCCCCGAGGCCCCUAGAGACGGCCAGGCCUACG UGCGGAAGGACGGCGAGUGGGUGCUGCUGAGCACCUUCC UG	
3′ UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCC CCUUGGGCCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACC CGUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGG C	4
Corresponding amino acid sequence	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVF RSSVLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPF NDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKV CEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVS QPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRD LPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWT AGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKC TLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATR FASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLN DLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTG CVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQ AGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFE LLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKK FLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTN TSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQT RAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPGSGGSVA SQSIIAYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTK TSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDK	11

	NTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDL LFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPL LTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRF NGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQ DVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDPPEAEVQ IDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLG QSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTT APAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNT FVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPD VDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYE QGSGYIPEAPRDGQAYVRKDGEWVLLSTFL	
PolyA tail	100 nt	
SARS-CoV-2 S Protein Variant 4		
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	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUA AGACCCCGGCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGUGAGCAGC CAGUGCGUGAACCUGACCACCCGGACCCAGCUGCCACCAG CCUACACCAACAGCUUCACCCGGGGCGUCUACUACCCCGA CAAGGUGUUCCGGAGCAGCGUCCUGCACAGCACCCAGGA CCUGUUCCUGCCCUUCUUCAGCAACGUGACCUGGUUCCAC GCCAUCCACGUGAGCGGCACCAACGGCACCAAGCGGUUCG ACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUCGC CAGCACCGAGAAGAGCAACAUCAUCCGGGGCUGGAUCUU CGGCACCACCCUGGACAGCAAGACCCAGAGCCUGCUGAUC GUGAAUAACGCCACCAACGUGGUGAUCAAGGUGUGCGAG UUCCAGUUCUGCAACGACCCCUUCCUGGGCGUGUACUACC ACAAGAACAACAAGAGCUGGAUGGAGAGCGAGUUCCGGG UGUACAGCAGCGCCAACAACUGCACCUUCGAGUACGUGA GCCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAGGGCA ACUUCAAGAACCUGCGGGAGUUCGUGUUCAAGAACAUCG ACGGCUACUUCAAGAUCUACAGCAAGCACACCCCAAUCA ACCUGGUGCGGGAUCUGCCCCAGGGCUUCUCAGCCCUGG AGCCCCUGGUGGACCUGCCCAUCGGCAUCAACAUCACCCG GUUCCAGACCCUGCUGGCCCUGCACCGGAGCUACCUGACC CCAGGCGACAGCAGCAGCGGGUGGACAGCAGGCGCGGCU GCUUACUACGUGGGCUACCUGCAGCCCCGGACCUUCCUGC UGAAGUACAACGAGAACGGCACCAUCACCGACGCCGUGG ACUGCGCCCUGGACCCUCUGAGCGAGACCAAGUGCACCCU GAAGAGCUUCACCGUGGAGAAGGGCAUCUACCAGACCAG CAACUUCCGGGUGCAGCCCACCGAGAGCAUCGUGCGGUU CCCCAACAUCACCAACCUGUGCCCCUUCGGCGAGGUGUUC AACGCCACCCGGUUCGCCAGCGUGUACGCCUGGAACCGGA AGCGGAUCAGCAACUGCGUGGCCGACUACAGCGUGCUGU ACAACAGCGCCAGCUUCAGCACCUUCAAGUGCUACGGCG UGAGCCCCACCAAGCUGAACGACCUGUGCUUCACCAACGU GUACGCCGACAGCUUCGUGAUCCGUGGCGACGAGGUGCG GCAGAUCGCACCCGGCCAGACAGGCAAGAUCGCCGACUAC AACUACAAGCUGCCCGACGACUUCACCGGCUGCGUGAUC GCCUGGAACAGCAACAACCUCGACAGCAAGGUGGGCGGC	13

	AACUACAACUACCUGUACCGGCUGUUCCGGAAGAGCAAC CUGAAGCCCUUCGAGCGGGACAUCAGCACCGAGAUCUAC CAAGCCGGCUCCACCCCUUGCAACGGCGUGGAGGGCUUCA ACUGCUACUUCCCUCUGCAGAGCUACGGCUUCCAGCCCAC CAACGGCGUGGGCUACCAGCCCUACCGGGUGGUGGUGCU GAGCUUCGAGCUGCUGCACGCCCCAGCCACCGUGUGUGGC CCCAAGAAGAGCACCAACCUGGUGAAGAACAAGUGCGUG AACUUCAACUUCAACGGCCUUACCGGCACCGGCGUGCUG ACCGAGAGCAACAAGAAAUUCCUGCCCUUUCAGCAGUUC GGCCGGGACAUCGCCGACACCACCGACGCUGUGCGGGAUC CCCAGACCCUGGAGAUCCUGGACAUCACCCCUUGCAGCUU CGGCGGCGUGAGCGUGAUCACCCCAGGCACCAACACCAGC AACCAGGUGGCCGUGCUGUACCAGGACGUGAACUGCACC GAGGUGCCCGUGGCCAUCCACGCCGACCAGCUGACACCCA CCUGGCGGGUCUACAGCACCGGCAGCAACGUGUUCCAGA CCCGGGCCGGUUGCCUGAUCGGCGCCGAGCACGUGAACA ACAGCUACGAGUGCGACAUCCCCAUCGGCGCCGGCAUCUG UGCCAGCUACCAGACCCAGACCAAUUCACCCCGGAGGGCA AGGAGCGUGGCCAGCCAGAGCAUCAUCGCCUACACCAUG AGCCUGGGCGCCGAGAACAGCGUGGCCUACAGCAACAAC AGCAUCGCCAUCCCCACCAACUUCACCAUCAGCGUGACCA CCGAGAUUCUGCCCGUGAGCAUGACCAAGACCAGCGUGG ACUGCACCAUGUACAUCUGCGGCGACAGCACCGAGUGCA GCAACCUGCUGCUGCAGUACGGCAGCUUCUGCACCCAGCU GAACCGGGCCCUGACCGGCAUCGCCGUGGAGCAGGACAA GAACACCCAGGAGGUGUUCGCCCAGGUGAAGCAGAUCUA CAAGACCCCUCCCAUCAAGGACUUCGGCGGCUUCAACUUC AGCCAGAUCCUGCCCGACCCCAGCAAGCCCAGCAAGCGGA GCUUCAUCGAGGACCUGCUGUUCAACAAGGUGACCCUAG CCGACGCCGGCUUCAUCAAGCAGUACGGCGACUGCCUCGG CGACAUAGCCGCCCGGGACCUGAUCUGCGCCCAGAAGUUC AACGGCCUGACCGUGCUGCCUCCCCUGCUGACCGACGAGA UGAUCGCCCAGUACACCAGCGCCCUGUUAGCCGGAACCAU CACCAGCGGCUGGACUUUCGGCGCUGGAGCCGCUCUGCA GAUCCCCUUCGCCAUGCAGAUGGCCUACCGGUUCAACGGC AUCGGCGUGACCCAGAACGUGCUGUACGAGAACCAGAAG CUGAUCGCCAACCAGUUCAACAGCGCCAUCGGCAAGAUCC AGGACAGCCUGAGCAGCACCGCUAGCGCCCUGGGCAAGC UGCAGGACGUGGUGAACCAGAACGCCCAGGCCCUGAACA CCCUGGUGAAGCAGCUGAGCAGCAACUUCGGCGCCAUCA GCAGCGUGCUGAACGACAUCCUGAGCCGGCUGGACCCUCC CGAGGCCGAGGUGCAGAUCGACCGGCUGAUCACUGGCCG GCUGCAGAGCCUGCAGACCUACGUGACCCAGCAGCUGAU CCGGGCCGCCGAGAUUCGGGCCAGCGCCAACCUGGCCGCC ACCAAGAUGAGCGAGUGCGUGCUGGGCCAGAGCAAGCGG GUGGACUUCUGCGGCAAGGGCUACCACCUGAUGAGCUUU CCCCAGAGCGCACCCCACGGAGUGGUGUUCCUGCACGUGA CCUACGUGCCCGCCCAGGAGAAGAACUUCACCACCGCCCC AGCCAUCUGCCACGACGGCAAGGCCCACUUUCCCCGGGAG GGCGUGUUCGUGAGCAACGGCACCCACUGGUUCGUGACC CAGCGGAACUUCUACGAGCCCCAGAUCAUCACCACCGACA ACACCUUCGUGAGCGGCAACUGCGACGUGGUGAUCGGCA UCGUGAACAACACCGUGUACGAUCCCCUGCAGCCCGAGCU GGACAGCUUCAAGGAGGAGCUGGACAAGUACUUCAAGAA UCACACCAGCCCCGACGUGGACCUGGGCGACAUCAGCGGC AUCAACGCCAGCGUGGUGAACAUCCAGAAGGAGAUCGAU CGGCUGAACGAGGUGGCCAAGAACCUGAACGAGAGCCUG AUCGACCUGCAGGAGCUGGGCAAGUACGAGCAGGGCAGC	

	GGCUACAUCCCCGAGGCCCCUAGAGACGGCCAGGCCUACG UGCGGAAGGACGGCGAGUGGGUGCUGCUGAGCACCUUCC UG	
3′ UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCC CCUUGGGCCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACC CGUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGG C	4
Corresponding amino acid sequence	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVF RSSVLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPF NDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKV CEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVS QPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRD LPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWT AGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKC TLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATR FASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLN DLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTG CVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQ AGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFE LLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKK FLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTN TSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQT RAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPRRARSVA SQSIIAYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTK TSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDK NTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDL LFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPL LTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRF NGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQ DVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDPPEAEVQ IDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLG QSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTT APAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNT FVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPD VDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYE QGSGYIPEAPRDGQAYVRKDGEWVLLSTFL	14
PolyA tail	100 nt	
SARS-CoV-2 S Protein Variant 5		
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	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUA AGACCCCGGCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGUGAGCAGC CAGUGCGUGAACCUGACCACCCGGACCCAGCUGCCACCAG CCUACACCAACAGCUUCACCCGGGGCGUCUACUACCCCGA CAAGGUGUUCCGGAGCAGCGUCCUGCACAGCACCCAGGA CCUGUUCCUGCCCUUCUUCAGCAACGUGACCUGGUUCCAC GCCAUCCACGUGAGCGGCACCAACGGCACCAAGCGGUUCG ACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUCGC CAGCACCGAGAAGAGCAACAUCAUCCGGGGCUGGAUCUU	16

	CGGCACCACCCUGGACAGCAAGACCCAGAGCCUGCUGAUC GUGAAUAACGCCACCAACGUGGUGAUCAAGGUGUGCGAG UUCCAGUUCUGCAACGACCCCUUCCUGGGCGUGUACUACC ACAAGAACAACAAGAGCUGGAUGGAGAGCGAGUUCCGGG UGUACAGCAGCGCCAACAACUGCACCUUCGAGUACGUGA GCCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAGGGCA ACUUCAAGAACCUGCGGGAGUUCGUGUUCAAGAACAUCG ACGGCUACUUCAAGAUCUACAGCAAGCACACCCCAAUCA ACCUGGUGCGGGAUCUGCCCCAGGGCUUCUCAGCCCUGG AGCCCCUGGUGGACCUGCCCAUCGGCAUCAACAUCACCCG GUUCCAGACCCUGCUGGCCCUGCACCGGAGCUACCUGACC CCAGGCGACAGCAGCAGCGGGUGGACAGCAGGCGCGGCU GCUUACUACGUGGGCUACCUGCAGCCCCGGACCUUCCUGC UGAAGUACAACGAGAACGGCACCAUCACCGACGCCGUGG ACUGCGCCCUGGACCCUCUGAGCGAGACCAAGUGCACCCU GAAGAGCUUCACCGUGGAGAAGGGCAUCUACCAGACCAG CAACUUCCGGGUGCAGCCCACCGAGAGCAUCGUGCGGUU CCCCAACAUCACCAACCUGUGCCCCUUCGGCGAGGUGUUC AACGCCACCCGGUUCGCCAGCGUGUACGCCUGGAACCGGA AGCGGAUCAGCAACUGCGUGGCCGACUACAGCGUGCUGU ACAACAGCGCCAGCUUCAGCACCUUCAAGUGCUACGGCG UGAGCCCCACCAAGCUGAACGACCUGUGCUUCACCAACGU GUACGCCGACAGCUUCGUGAUCCGUGGCGACGAGGUGCG GCAGAUCGCACCCGGCCAGACAGGCAAGAUCGCCGACUAC AACUACAAGCUGCCCGACGACUUCACCGGCUGCGUGAUC GCCUGGAACAGCAACAACCUCGACAGCAAGGUGGGCGGC AACUACAACUACCUGUACCGGCUGUUCCGGAAGAGCAAC CUGAAGCCCUUCGAGCGGGACAUCAGCACCGAGAUCUAC CAAGCCGGCUCCACCCCUUGCAACGGCGUGGAGGGCUUCA ACUGCUACUUCCCUCUGCAGAGCUACGGCUUCCAGCCCAC CAACGGCGUGGGCUACCAGCCCUACCGGGUGGUGGUGCU GAGCUUCGAGCUGCUGCACGCCCCAGCCACCGUGUGUGGC CCCAAGAAGAGCACCAACCUGGUGAAGAACAAGUGCGUG AACUUCAACUUCAACGGCCUUACCGGCACCGGCGUGCUG ACCGAGAGCAACAAGAAAUUCCUGCCCUUUCAGCAGUUC GGCCGGGACAUCGCCGACACCACCGACGCUGUGCGGGAUC CCCAGACCCUGGAGAUCCUGGACAUCACCCCUUGCAGCUU CGGCGGCGUGAGCGUGAUCACCCCAGGCACCAACACCAGC AACCAGGUGGCCGUGCUGUACCAGGACGUGAACUGCACC GAGGUGCCCGUGGCCAUCCACGCCGACCAGCUGACACCCA CCUGGCGGGUCUACAGCACCGGCAGCAACGUGUUCCAGA CCCGGGCCGGUUGCCUGAUCGGCGCCGAGCACGUGAACA ACAGCUACGAGUGCGACAUCCCCAUCGGCGCCGGCAUCUG UGCCAGCUACCAGACCCAGACCAAUUCACCCGGCAGCGGC GGCAGCGUGGCCAGCCAGAGCAUCAUCGCCUACACCAUG AGCCUGGGCGCCGAGAACAGCGUGGCCUACAGCAACAAC AGCAUCGCCAUCCCCACCAACUUCACCAUCAGCGUGACCA CCGAGAUUCUGCCCGUGAGCAUGACCAAGACCAGCGUGG ACUGCACCAUGUACAUCUGCGGCGACAGCACCGAGUGCA GCAACCUGCUGCUGCAGUACGGCAGCUUCUGCACCCAGCU GAACCGGGCCCUGACCGGCAUCGCCGUGGAGCAGGACAA GAACACCCAGGAGGUGUUCGCCCAGGUGAAGCAGAUCUA CAAGACCCCUCCCAUCAAGGACUUCGGCGGCUUCAACUUC AGCCAGAUCCUGCCCGACCCCAGCAAGCCCAGCAAGCGGA GCUUCAUCGAGGACCUGCUGUUCAACAAGGUGACCCUAG CCGACGCCGGCUUCAUCAAGCAGUACGGCGACUGCCUCGG CGACAUAGCCGCCCGGGACCUGAUCUGCGCCCAGAAGUUC AACGGCCUGACCGUGCUGCCUCCCCUGCUGACCGACGAGA	

	UGAUCGCCCAGUACACCAGCGCCCUGUUAGCCGGAACCAU CACCAGCGGCUGGACUUUCGGCGCUGGAGCCGCUCUGCA GAUCCCCUUCGCCAUGCAGAUGGCCUACCGGUUCAACGGC AUCGGCGUGACCCAGAACGUGCUGUACGAGAACCAGAAG CUGAUCGCCAACCAGUUCAACAGCGCCAUCGGCAAGAUCC AGGACAGCCUGAGCAGCACCGCUAGCGCCCUGGGCAAGC UGCAGGACGUGGUGAACCAGAACGCCCAGGCCCUGAACA CCCUGGUGAAGCAGCUGAGCAGCAACUUCGGCGCCAUCA GCAGCGUGCUGAACGACAUCCUGAGCCGGCUGGACCCUCC CGAGGCCGAGGUGCAGAUCGACCGGCUGAUCACUGGCCG GCUGCAGAGCCUGCAGACCUACGUGACCCAGCAGCUGAU CCGGGCCGCCGAGAUUCGGGCCAGCGCCAACCUGGCCGCC ACCAAGAUGAGCGAGUGCGUGCUGGGCCAGAGCAAGCGG GUGGACUUCUGCGGCAAGGGCUACCACCUGAUGAGCUUU CCCCAGAGCGCACCCCACGGAGUGGUGUUCCUGCACGUGA CCUACGUGCCCGCCCAGGAGAAGAACUUCACCACCGCCCC AGCCAUCUGCCACGACGGCAAGGCCCACUUUCCCCGGGAG GGCGUGUUCGUGAGCAACGGCACCCACUGGUUCGUGACC CAGCGGAACUUCUACGAGCCCCAGAUCAUCACCACCGACA ACACCUUCGUGAGCGGCAACUGCGACGUGGUGAUCGGCA UCGUGAACAACACCGUGUACGAUCCCCUGCAGCCCGAGCU GGACAGCUUCAAGGAGGAGCUGGACAAGUACUUCAAGAA UCACACCAGCCCCGACGUGGACCUGGGCGACAUCAGCGGC AUCAACGCCAGCGUGGUGAACAUCCAGAAGGAGAUCGAU CGGCUGAACGAGGUGGCCAAGAACCUGAACGAGAGCCUG AUCGACCUGCAGGAGCUGGGCAAGUACGAGCAGUACAUC AAGUGGCCCUGGUACAUCUGGCUGGGCUUCAUCGCCGGC CUGAUCGCCAUCGUGAUGGUGACCAUCAUGCUGUGCUGC AUGACCAGCUGCUGCAGCUGCCUGAAGGGCUGUUGCAGC UGCGGCAGCUGCUGCAAGUUCGACGAGGACGACAGCGAG CCCGUGCUGAAGGGCGUGAAGCUGCACUACACC	
3′ UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCC CCUUGGGCCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACC CGUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGG C	4
Corresponding amino acid sequence	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVF RSSVLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPF NDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKV CEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVS QPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRD LPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWT AGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKC TLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATR FASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLN DLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTG CVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQ AGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFE LLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKK FLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTN TSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQT RAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPGSGGSVA SQSIIAYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTK TSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDK NTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDL LFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPL LTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRF NGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQ	17

	DVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDPPEAEVQ IDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLG QSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTT APAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNT FVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPD VDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYE QYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCLKGCCSC GSCCKFDEDDSEPVLKGVKLHYT	
PolyA tail	100 nt	
SARS-CoV-2 S Protein Variant 6		
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	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUA AGACCCCGGCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGUGAGCAGC CAGUGCGUGAACCUGACCACCCGGACCCAGCUGCCACCAG CCUACACCAACAGCUUCACCCGGGGCGUCUACUACCCCGA CAAGGUGUUCCGGAGCAGCGUCCUGCACAGCACCCAGGA CCUGUUCCUGCCCUUCUUCAGCAACGUGACCUGGUUCCAC GCCAUCCACGUGAGCGGCACCAACGGCACCAAGCGGUUCG ACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUCGC CAGCACCGAGAAGAGCAACAUCAUCCGGGGCUGGAUCUU CGGCACCACCCUGGACAGCAAGACCCAGAGCCUGCUGAUC GUGAAUAACGCCACCAACGUGGUGAUCAAGGUGUGCGAG UUCCAGUUCUGCAACGACCCCUUCCUGGGCGUGUACUACC ACAAGAACAACAAGAGCUGGAUGGAGAGCGAGUUCCGGG UGUACAGCAGCGCCAACAACUGCACCUUCGAGUACGUGA GCCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAGGGCA ACUUCAAGAACCUGCGGGAGUUCGUGUUCAAGAACAUCG ACGGCUACUUCAAGAUCUACAGCAAGCACACCCCAAUCA ACCUGGUGCGGGAUCUGCCCCAGGGCUUCUCAGCCCUGG AGCCCCUGGUGGACCUGCCCAUCGGCAUCAACAUCACCCG GUUCCAGACCCUGCUGGCCCUGCACCGGAGCUACCUGACC CCAGGCGACAGCAGCAGCGGGUGGACAGCAGGCGCGGCU GCUUACUACGUGGGCUACCUGCAGCCCCGGACCUUCCUGC UGAAGUACAACGAGAACGGCACCAUCACCGACGCCGUGG ACUGCGCCCUGGACCCUCUGAGCGAGACCAAGUGCACCCU GAAGAGCUUCACCGUGGAGAAGGGCAUCUACCAGACCAG CAACUUCCGGGUGCAGCCCACCGAGAGCAUCGUGCGGUU CCCCAACAUCACCAACCUGUGCCCCUUCGGCGAGGUGUUC AACGCCACCCGGUUCGCCAGCGUGUACGCCUGGAACCGGA AGCGGAUCAGCAACUGCGUGGCCGACUACAGCGUGCUGU ACAACAGCGCCAGCUUCAGCACCUUCAAGUGCUACGGCG UGAGCCCCACCAAGCUGAACGACCUGUGCUUCACCAACGU GUACGCCGACAGCUUCGUGAUCCGUGGCGACGAGGUGCG GCAGAUCGCACCCGGCCAGACAGGCAAGAUCGCCGACUAC AACUACAAGCUGCCCGACGACUUCACCGGCUGCGUGAUC GCCUGGAACAGCAACAACCUCGACAGCAAGGUGGGCGGC AACUACAACUACCUGUACCGGCUGUUCCGGAAGAGCAAC CUGAAGCCCUUCGAGCGGGACAUCAGCACCGAGAUCUAC CAAGCCGGCUCCACCCCUUGCAACGGCGUGGAGGGCUUCA	19

	ACUGCUACUUCCCUCUGCAGAGCUACGGCUUCCAGCCCAC CAACGGCGUGGGCUACCAGCCCUACCGGGUGGUGGUGCU GAGCUUCGAGCUGCUGCACGCCCCAGCCACCGUGUGUGGC CCCAAGAAGAGCACCAACCUGGUGAAGAACAAGUGCGUG AACUUCAACUUCAACGGCCUUACCGGCACCGGCGUGCUG ACCGAGAGCAACAAGAAAUUCCUGCCCUUUCAGCAGUUC GGCCGGGACAUCGCCGACACCACCGACGCUGUGCGGGAUC CCCAGACCCUGGAGAUCCUGGACAUCACCCCUUGCAGCUU CGGCGGCGUGAGCGUGAUCACCCCAGGCACCAACACCAGC AACCAGGUGGCCGUGCUGUACCAGGACGUGAACUGCACC GAGGUGCCCGUGGCCAUCCACGCCGACCAGCUGACACCCA CCUGGCGGGUCUACAGCACCGGCAGCAACGUGUUCCAGA CCCGGGCCGGUUGCCUGAUCGGCGCCGAGCACGUGAACA ACAGCUACGAGUGCGACAUCCCCAUCGGCGCCGGCAUCUG UGCCAGCUACCAGACCCAGACCAAUUCACCCCGGAGGGCA AGGAGCGUGGCCAGCCAGAGCAUCAUCGCCUACACCAUG AGCCUGGGCGCCGAGAACAGCGUGGCCUACAGCAACAAC AGCAUCGCCAUCCCCCACCAACUUCACCAUCAGCGUGACCA CCGAGAUUCUGCCCGUGAGCAUGACCAAGACCAGCGUGG ACUGCACCAUGUACAUCUGCGGCGACAGCACCGAGUGCA GCAACCUGCUGCUGCAGUACGGCAGCUUCUGCACCCAGCU GAACCGGGCCCUGACCGGCAUCGCCGUGGAGCAGGACAA GAACACCCAGGAGGUGUUCGCCCAGGUGAAGCAGAUCUA CAAGACCCCUCCCAUCAAGGACUUCGGCGGCUUCAACUUC AGCCAGAUCCUGCCCGACCCCAGCAAGCCCAGCAAGCGGA GCUUCAUCGAGGACCUGCUGUUCAACAAGGUGACCCUAG CCGACGCCGGCUUCAUCAAGCAGUACGGCGACUGCCUCGG CGACAUAGCCGCCCGGGACCUGAUCUGCGCCCAGAAGUUC AACGGCCUGACCGUGCUGCCUCCCCUGCUGACCGACGAGA UGAUCGCCCAGUACACCAGCGCCCUGUUAGCCGGAACCAU CACCAGCGGCUGGACUUUCGGCGCUGGAGCCGCUCUGCA GAUCCCCUUCGCCAUGCAGAUGGCCUACCGGUUCAACGGC AUCGGCGUGACCCAGAACGUGCUGUACGAGAACCAGAAG CUGAUCGCCAACCAGUUCAACAGCGCCAUCGGCAAGAUCC AGGACAGCCUGAGCAGCACCGCUAGCGCCCUGGGCAAGC UGCAGGACGUGGUGAACCAGAACGCCCAGGCCCUGAACA CCCUGGUGAAGCAGCUGAGCAGCAACUUCGGCGCCAUCA GCAGCGUGCUGAACGACAUCCUGAGCCGGCUGGACCCUCC CGAGGCCGAGGUGCAGAUCGACCGGCUGAUCACUGGCCG GCUGCAGAGCCUGCAGACCUACGUGACCCAGCAGCUGAU CCGGGCCGCCGAGAUUCGGGCCAGCGCCAACCUGGCCGCC ACCAAGAUGAGCGAGUGCGUGCUGGGCCAGAGCAAGCGG GUGGACUUCUGCGGCAAGGGCUACCACCUGAUGAGCUUU CCCCAGAGCGCACCCCACGGAGUGGUGUUCCUGCACGUGA CCUACGUGCCCGCCCAGGAGAAGAACUUCACCACCGCCCC AGCCAUCUGCCACGACGGCAAGGCCCACUUUCCCCGGGAG GGCGUGUUCGUGAGCAACGGCACCCACUGGUUCGUGACC CAGCGGAACUUCUACGAGCCCCAGAUCAUCACCACCGACA ACACCUUCGUGAGCGGCAACUGCGACGUGGUGAUCGGCA UCGUGAACAACACCGUGUACGAUCCCCUGCAGCCCGAGCU GGACAGCUUCAAGGAGGAGCUGGACAAGUACUUCAAGAA UCACACCAGCCCCGACGUGGACCUGGGCGACAUCAGCGGC AUCAACGCCAGCGUGGUGAACAUCCAGAAGGAGAUCGAU CGGCUGAACGAGGUGGCCAAGAACCUGAACGAGAGCCUG AUCGACCUGCAGGAGCUGGGCAAGUACGAGCAGUACAUC AAGUGGCCCUGGUACAUCUGGCUGGGCUUCAUCGCCGGC CUGAUCGCCAUCGUGAUGGUGACCAUCAUGCUG	

3′ UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCC CCUUGGGCCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACC CGUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGG C	4
Corresponding amino acid sequence	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVF RSSVLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPF NDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKV CEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVS QPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRD LPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWT AGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKC TLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATR FASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLN DLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTG CVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQ AGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFE LLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKK FLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTN TSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQT RAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPRRARSVA SQSIIAYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTK TSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDK NTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDL LFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPL LTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRF NGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQ DVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDPPEAEVQ IDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLG QSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTT APAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNT FVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPD VDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYE QYIKWPWYIWLGFIAGLIAIVMVTIML	20
PolyA tail	100 nt	
SARS-CoV-2 S Protein Variant 7		
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	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUA AGACCCCGGCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGUGAGCAGC CAGUGCGUGAACCUGACCACCCGGACCCAGCUGCCACCAG CCUACACCAACAGCUUCACCCGGGGCGUCUACUACCCCGA CAAGGUGUUCCGGAGCAGCGUCCUGCACAGCACCCAGGA CCUGUUCCUGCCCUUCUUCAGCAACGUGACCUGGUUCCAC GCCAUCCACGUGAGCGGCACCAACGGCACCAAGCGGUUCG ACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUCGC CAGCACCGAGAAGAGCAACAUCAUCCGGGGCUGGAUCUU CGGCACCACCCUGGACAGCAAGACCCAGAGCCUGCUGAUC GUGAAUAACGCCACCAACGUGGUGAUCAAGGUGUGCGAG UUCCAGUUCUGCAACGACCCCUUCCUGGGCGUGUACUACC	22

	ACAAGAACAACAAGAGCUGGAUGGAGAGCGAGUUCCGGG UGUACAGCAGCGCCAACAACUGCACCUUCGAGUACGUGA GCCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAGGGCA ACUUCAAGAACCUGCGGGAGUUCGUGUUCAAGAACAUCG ACGGCUACUUCAAGAUCUACAGCAAGCACACCCCAAUCA ACCUGGUGCGGGAUCUGCCCCAGGGCUUCUCAGCCCUGG AGCCCCUGGUGGACCUGCCCAUCGGCAUCAACAUCACCCG GUUCCAGACCCUGCUGGCCCUGCACCGGAGCUACCUGACC CCAGGCGACAGCAGCAGCGGGUGGACAGCAGGCGCGGCU GCUUACUACGUGGGCUACCUGCAGCCCCGGACCUUCCUGC UGAAGUACAACGAGAACGGCACCAUCACCGACGCCGUGG ACUGCGCCCUGGACCCUCUGAGCGAGACCAAGUGCACCCU GAAGAGCUUCACCGUGGAGAAGGGCAUCUACCAGACCAG CAACUUCCGGGUGCAGCCCACCGAGAGCAUCGUGCGGUU CCCCAACAUCACCAACCUGUGCCCCUUCGGCGAGGUGUUC AACGCCACCCGGUUCGCCAGCGUGUACGCCUGGAACCGGA AGCGGAUCAGCAACUGCGUGGCCGACUACAGCGUGCUGU ACAACAGCGCCAGCUUCAGCACCUUCAAGUGCUACGGCG UGAGCCCCACCAAGCUGAACGACCUGUGCUUCACCAACGU GUACGCCGACAGCUUCGUGAUCCGUGGCGACGAGGUGCG GCAGAUCGCACCCGGCCAGACAGGCAAGAUCGCCGACUAC AACUACAAGCUGCCCGACGACUUCACCGGCUGCGUGAUC GCCUGGAACAGCAACAACCUCGACAGCAAGGUGGGCGGC AACUACAACUACCUGUACCGGCUGUUCCGGAAGAGCAAC CUGAAGCCCUUCGAGCGGGACAUCAGCACCGAGAUCUAC CAAGCCGGCUCCACCCCUUGCAACGGCGUGGAGGGCUUCA ACUGCUACUUCCCUCUGCAGAGCUACGGCUUCCAGCCCAC CAACGGCGUGGGCUACCAGCCCUACCGGGUGGUGGUGCU GAGCUUCGAGCUGCUGCACGCCCCAGCCACCGUGUGUGGC CCCAAGAAGAGCACCAACCUGGUGAAGAACAAGUGCGUG AACUUCAACUUCAACGGCCUUACCGGCACCGGCGUGCUG ACCGAGAGCAACAAGAAAUUCCUGCCCUUUCAGCAGUUC GGCCGGGACAUCGCCGACACCACCGACGCUGUGCGGGAUC CCCAGACCCUGGAGAUCCUGGACAUCACCCCUUGCAGCUU CGGCGGCGUGAGCGUGAUCACCCCAGGCACCAACACCAGC AACCAGGUGGCCGUGCUGUACCAGGACGUGAACUGCACC GAGGUGCCCGUGGCCAUCCACGCCGACCAGCUGACACCCA CCUGGCGGGUCUACAGCACCGGCAGCAACGUGUUCCAGA CCCGGGCCGGUUGCCUGAUCGGCGCCGAGCACGUGAACA ACAGCUACGAGUGCGACAUCCCCAUCGGCGCCGGCAUCUG UGCCAGCUACCAGACCCAGACCGUGUCACUGAGGAGCGU GGCCAGCCAGAGCAUCAUCGCCUACACCAUGAGCCUGGGC GCCGAGAACAGCGUGGCCUACAGCAACAACAGCAUCGCC AUCCCCACCAACUUCACCAUCAGCGUGACCACCGAGAUUC UGCCCGUGAGCAUGACCAAGACCAGCGUGGACUGCACCA UGUACAUCUGCGGCGACAGCACCGAGUGCAGCAACCUGC UGCUGCAGUACGGCAGCUUCUGCACCCAGCUGAACCGGG CCCUGACCGGCAUCGCCGUGGAGCAGGACAAGAACACCCA GGAGGUGUUCGCCCAGGUGAAGCAGAUCUACAAGACCCC UCCCAUCAAGGACUUCGGCGGCUUCAACUUCAGCCAGAU CCUGCCCGACCCCAGCAAGCCCAGCAAGCGGAGCUUCAUC GAGGACCUGCUGUUCAACAAGGUGACCCUAGCCGACGCC GGCUUCAUCAAGCAGUACGGCGACUGCCUCGGCGACAUA GCCGCCCGGGACCUGAUCUGCGCCCAGAAGUUCAACGGCC UGACCGUGCUGCCUCCCCUGCUGACCGACGAGAUGAUCGC CCAGUACACCAGCGCCCUGUUAGCCGGAACCAUCACCAGC GGCUGGACUUUCGGCGCUGGAGCCGCUCUGCAGAUCCCC UUCGCCAUGCAGAUGGCCUACCGGUUCAACGGCAUCGGC	

	GUGACCCAGAACGUGCUGUACGAGAACCAGAAGCUGAUC GCCAACCAGUUCAACAGCGCCAUCGGCAAGAUCCAGGAC AGCCUGAGCAGCACCGCUAGCGCCCUGGGCAAGCUGCAG GACGUGGUGAACCAGAACGCCCAGGCCCUGAACACCCUG GUGAAGCAGCUGAGCAGCAACUUCGGCGCCAUCAGCAGC GUGCUGAACGACAUCCUGAGCCGGCUGGACAAGGUGGAG GCCGAGGUGCAGAUCGACCGGCUGAUCACUGGCCGGCUG CAGAGCCUGCAGACCUACGUGACCCAGCAGCUGAUCCGG GCCGCCGAGAUUCGGGCCAGCGCCAACCUGGCCGCCACCA AGAUGAGCGAGUGCGUGCUGGGCCAGAGCAAGCGGGUGG ACUUCUGCGGCAAGGGCUACCACCUGAUGAGCUUUCCCC AGAGCGCACCCCACGGAGUGGUGUUCCUGCACGUGACCU ACGUGCCCGCCCAGGAGAAGAACUUCACCACCGCCCCAGC CAUCUGCCACGACGGCAAGGCCCACUUUCCCCGGGAGGGC GUGUUCGUGAGCAACGGCACCCACUGGUUCGUGACCCAG CGGAACUUCUACGAGCCCCAGAUCAUCACCACCGACAACA CCUUCGUGAGCGGCAACUGCGACGUGGUGAUCGGCAUCG UGAACAACACCGUGUACGAUCCCCUGCAGCCCGAGCUGG ACAGCUUCAAGGAGGAGCUGGACAAGUACUUCAAGAAUC ACACCAGCCCCGACGUGGACCUGGGCGACAUCAGCGGCAU CAACGCCAGCGUGGUGAACAUCCAGAAGGAGAUCGAUCG GCUGAACGAGGUGGCCAAGAACCUGAACGAGAGCCUGAU CGACCUGCAGGAGCUGGGCAAGUACGAGCAGUACAUCAA GUGGCCCUGGUACAUCUGGCUGGGCUUCAUCGCCGGCCU GAUCGCCAUCGUGAUGGUGACCAUCAUGCUGUGCUGCAU GACCAGCUGCUGCAGCUGCCUGAAGGGCUGUUGCAGCUG CGGCAGCUGCUGCAAGUUCGACGAGGACGACAGCGAGCC CGUGCUGAAGGGCGUGAAGCUGCACUACACC	
3′ UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCC CCUUGGGCCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACC CGUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGG C	4
Corresponding amino acid sequence	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVF RSSVLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPF NDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKV CEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVS QPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRD LPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWT AGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKC TLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATR FASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLN DLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTG CVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQ AGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFE LLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKK FLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTN TSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQT RAGCLIGAEHVNNSYECDIPIGAGICASYQTQTVSLRSVASQSII AYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVD CTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQE VFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNK VTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPLLTDE MIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIG VTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVN QNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLI TGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKR VDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAI	23

	CHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFVSG NCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLG DISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIK WPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCLKGCCSCGSCC KFDEDDSEPVLKGVKLHYT	
PolyA tail	100 nt	
SARS-CoV-2 S Protein Variant 8		
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	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUA AGACCCCGGCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGUGAGCAGC CAGUGCGUGAACCUGACCACCCGGACCCAGCUGCCACCAG CCUACACCAACAGCUUCACCCGGGGCGUCUACUACCCCGA CAAGGUGUUCCGGAGCAGCGUCCUGCACAGCACCCAGGA CCUGUUCCUGCCCUUCUUCAGCAACGUGACCUGGUUCCAC GCCAUCCACGUGAGCGGCACCAACGGCACCAAGCGGUUCG ACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUCGC CAGCACCGAGAAGAGCAACAUCAUCCGGGGCUGGAUCUU CGGCACCACCCUGGACAGCAAGACCCAGAGCCUGCUGAUC GUGAAUAACGCCACCAACGUGGUGAUCAAGGUGUGCGAG UUCCAGUUCUGCAACGACCCCUUCCUGGGCGUGUACUACC ACAAGAACAACAAGAGCUGGAUGGAGAGCGAGUUCCGGG UGUACAGCAGCGCCAACAACUGCACCUUCGAGUACGUGA GCCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAGGGCA ACUUCAAGAACCUGCGGGAGUUCGUGUUCAAGAACAUCG ACGGCUACUUCAAGAUCUACAGCAAGCACACCCCAAUCA ACCUGGUGCGGGAUCUGCCCCAGGGCUUCUCAGCCCUGG AGCCCCUGGUGGACCUGCCCAUCGGCAUCAACAUCACCCG GUUCCAGACCCUGCUGGCCCUGCACCGGAGCUACCUGACC CCAGGCGACAGCAGCAGCGGGUGGACAGCAGGCGCGGCU GCUUACUACGUGGGCUACCUGCAGCCCCGGACCUUCCUGC UGAAGUACAACGAGAACGGCACCAUCACCGACGCCGUGG ACUGCGCCCUGGACCCUCUGAGCGAGACCAAGUGCACCCU GAAGAGCUUCACCGUGGAGAAGGGCAUCUACCAGACCAG CAACUUCCGGGUGCAGCCCACCGAGAGCAUCGUGCGGUU CCCCAACAUCACCAACCUGUGCCCCUUCGGCGAGGUGUUC AACGCCACCCGGUUCGCCAGCGUGUACGCCUGGAACCGGA AGCGGAUCAGCAACUGCGUGGCCGACUACAGCGUGCUGU ACAACAGCGCCAGCUUCAGCACCUUCAAGUGCUACGGCG UGAGCCCCACCAAGCUGAACGACCUGUGCUUCACCAACGU GUACGCCGACAGCUUCGUGAUCCGUGGCGACGAGGUGCG GCAGAUCGCACCCGGCCAGACAGGCAAGAUCGCCGACUAC AACUACAAGCUGCCCGACGACUUCACCGGCUGCGUGAUC GCCUGGAACAGCAACAACCUCGACAGCAAGGUGGGCGGC AACUACAACUACCUGUACCGGCUGUUCCGGAAGAGCAAC CUGAAGCCCUUCGAGCGGGACAUCAGCACCGAGAUCUAC CAAGCCGGCUCCACCCCUUGCAACGGCGUGGAGGGCUUCA ACUGCUACUUCCCUCUGCAGAGCUACGGCUUCCAGCCCAC CAACGGCGUGGGCUACCAGCCCUACCGGGUGGUGGUGCU GAGCUUCGAGCUGCUGCACGCCCCAGCCACCGUGUGUGGC	25

	CCCAAGAAGAGCACCAACCUGGUGAAGAACAAGUGCGUG AACUUCAACUUCAACGGCCUUACCGGCACCGGCGUGCUG ACCGAGAGCAACAAGAAAUUCCUGCCCUUUCAGCAGUUC GGCCGGGACAUCGCCGACACCACCGACGCUGUGCGGGAUC CCCAGACCCUGGAGAUCCUGGACAUCACCCCUUGCAGCUU CGGCGGCGUGAGCGUGAUCACCCCAGGCACCAACACCAGC AACCAGGUGGCCGUGCUGUACCAGGACGUGAACUGCACC GAGGUGCCCGUGGCCAUCCACGCCGACCAGCUGACACCCA CCUGGCGGGUCUACAGCACCGGCAGCAACGUGUUCCAGA CCCGGGCCGGUUGCCUGAUCGGCGCCGAGCACGUGAACA ACAGCUACGAGUGCGACAUCCCCAUCGGCGCCGGCAUCUG UGCCAGCUACCAGACCCAGACCAAUUCACCCCGGAGGGCA AGGAGCGUGGCCAGCCAGAGCAUCAUCGCCUACACCAUG AGCCUGGGCGCCGAGAACAGCGUGGCCUACAGCAACAAC AGCAUCGCCAUCCCCACCAACUUCACCAUCAGCGUGACCA CCGAGAUUCUGCCCGUGAGCAUGACCAAGACCAGCGUGG ACUGCACCAUGUACAUCUGCGGCGACAGCACCGAGUGCA GCAACCUGCUGCUGCAGUACGGCAGCUUCUGCACCCAGCU GAACCGGGCCCUGACCGGCAUCGCCGUGGAGCAGGACAA GAACACCCAGGAGGUGUUCGCCCAGGUGAAGCAGAUCUA CAAGACCCCUCCCAUCAAGGACUUCGGCGGCUUCAACUUC AGCCAGAUCCUGCCCGACCCCAGCAAGCCCAGCAAGCGGA GCUUCAUCGAGGACCUGCUGUUCAACAAGGUGACCCUAG CCGACGCCGGCUUCAUCAAGCAGUACGGCGACUGCCUCGG CGACAUAGCCGCCCGGGACCUGAUCUGCGCCCAGAAGUUC AACGGCCUGACCGUGCUGCCUCCCCUGCUGACCGACGAGA UGAUCGCCCAGUACACCAGCGCCCUGUUAGCCGGAACCAU CACCAGCGGCUGGACUUUCGGCGCUGGAGCCGCUCUGCA GAUCCCCUUCGCCAUGCAGAUGGCCUACCGGUUCAACGGC AUCGGCGUGACCCAGAACGUGCUGUACGAGAACCAGAAG CUGAUCGCCAACCAGUUCAACAGCGCCAUCGGCAAGAUCC AGGACAGCCUGAGCAGCACCGCUAGCGCCCUGGGCAAGC UGCAGGACGUGGUGAACCAGAACGCCCAGGCCCUGAACA CCCUGGUGAAGCAGCUGAGCAGCAACUUCGGCGCCAUCA GCAGCGUGCUGAACGACAUCCUGAGCCGGCUGGACAAGG UGGAGGCCGAGGUGCAGAUCGACCGGCUGAUCACUGGCC GGCUGCAGAGCCUGCAGACCUACGUGACCCAGCAGCUGA UCCGGGCCGCCGAGAUUCGGGCCAGCGCCAACCUGGCCGC CACCAAGAUGAGCGAGUGCGUGCUGGGCCAGAGCAAGCG GGUGGACUUCUGCGGCAAGGGCUACCACCUGAUGAGCUU UCCCCAGAGCGCACCCCACGGAGUGGUGUUCCUGCACGUG ACCUACGUGCCCGCCCAGGAGAAGAACUUCACCACCGCCC CAGCCAUCUGCCACGACGGCAAGGCCCACUUUCCCCGGGA GGGCGUGUUCGUGAGCAACGGCACCCACUGGUUCGUGAC CCAGCGGAACUUCUACGAGCCCCAGAUCAUCACCACCGAC AACACCUUCGUGAGCGGCAACUGCGACGUGGUGAUCGGC AUCGUGAACAACACCGUGUACGAUCCCCUGCAGCCCGAGC UGGACAGCUUCAAGGAGGAGCUGGACAAGUACUUCAAGA AUCACACCAGCCCCGACGUGGACCUGGGCGACAUCAGCGG CAUCAACGCCAGCGUGGUGAACAUCCAGAAGGAGAUCGA UCGGCUGAACGAGGUGGCCAAGAACCUGAACGAGAGCCU GAUCGACCUGCAGGAGCUGGGCAAGUACGAGCAGUACAU CAAGUGGCCCUGGUACAUCUGGCUGGGCUUCAUCGCCGG CCUGAUCGCCAUCGUGAUGGUGACCAUCAUGCUGUGCUG CAUGACCAGCUGCUGCAGCUGCCUGAAGGGCUGUUGCAG CUGCGGCAGCUGCUGCAAGUUCGACGAGGACGACAGCGA GCCCGUGCUGAAGGGCGUG	

3′ UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCC CCUUGGGCCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACC CGUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGG C	4
Corresponding amino acid sequence	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVF RSSVLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPF NDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKV CEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVS QPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRD LPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWT AGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKC TLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATR FASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLN DLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTG CVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQ AGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFE LLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKK FLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTN TSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQT RAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPRRARSVA SQSIIAYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTK TSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDK NTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDL LFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPL LTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRF NGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQ DVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEV QIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVL GQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNF TTAPAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTD NTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTS PDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGK YEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCLKGCC SCGSCCKFDEDDSEPVLKGV	26
PolyA tail	100 nt	
SARS-CoV-2 S Protein Variant 9		
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; SEQ ID NO: 105 corresponds to fully modified RNA sequence including 5′ UTR, mRNA ORF, and 3′UTR; SEQ ID NO: 106 corresponds to fully modified ORF of mRNA construct.	105 106
Cap	7mG(5′)ppp(5′)NlmpNp	
5′ UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUA AGACCCCGGCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGUGAGCAGC CAGUGCGUGAACCUGACCACCCGGACCCAGCUGCCACCAG CCUACACCAACAGCUUCACCCGGGGCGUCUACUACCCCGA CAAGGUGUUCCGGAGCAGCGUCCUGCACAGCACCCAGGA CCUGUUCCUGCCCUUCUUCAGCAACGUGACCUGGUUCCAC GCCAUCCACGUGAGCGGCACCAACGGCACCAAGCGGUUCG ACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUCGC	28

	CAGCACCGAGAAGAGCAACAUCAUCCGGGGCUGGAUCUU CGGCACCACCCUGGACAGCAAGACCCAGAGCCUGCUGAUC GUGAAUAACGCCACCAACGUGGUGAUCAAGGUGUGCGAG UUCCAGUUCUGCAACGACCCCUUCCUGGGCGUGUACUACC ACAAGAACAACAAGAGCUGGAUGGAGAGCGAGUUCCGGG UGUACAGCAGCGCCAACAACUGCACCUUCGAGUACGUGA GCCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAGGGCA ACUUCAAGAACCUGCGGGAGUUCGUGUUCAAGAACAUCG ACGGCUACUUCAAGAUCUACAGCAAGCACACCCCAAUCA ACCUGGUGCGGGAUCUGCCCCAGGGCUUCUCAGCCCUGG AGCCCCUGGUGGACCUGCCCAUCGGCAUCAACAUCACCCG GUUCCAGACCCUGCUGGCCCUGCACCGGAGCUACCUGACC CCAGGCGACAGCAGCAGCGGGUGGACAGCAGGCGCGGCU GCUUACUACGUGGGCUACCUGCAGCCCCGGACCUUCCUGC UGAAGUACAACGAGAACGGCACCAUCACCGACGCCGUGG ACUGCGCCCUGGACCCUCUGAGCGAGACCAAGUGCACCCU GAAGAGCUUCACCGUGGAGAAGGGCAUCUACCAGACCAG CAACUUCCGGGUGCAGCCCACCGAGAGCAUCGUGCGGUU CCCCAACAUCACCAACCUGUGCCCCUUCGGCGAGGUGUUC AACGCCACCCGGUUCGCCAGCGUGUACGCCUGGAACCGGA AGCGGAUCAGCAACUGCGUGGCCGACUACAGCGUGCUGU ACAACAGCGCCAGCUUCAGCACCUUCAAGUGCUACGGCG UGAGCCCCACCAAGCUGAACGACCUGUGCUUCACCAACGU GUACGCCGACAGCUUCGUGAUCCGUGGCGACGAGGUGCG GCAGAUCGCACCCGGCCAGACAGGCAAGAUCGCCGACUAC AACUACAAGCUGCCCGACGACUUCACCGGCUGCGUGAUC GCCUGGAACAGCAACAACCUCGACAGCAAGGUGGGCGGC AACUACAACUACCUGUACCGGCUGUUCCGGAAGAGCAAC CUGAAGCCCUUCGAGCGGGACAUCAGCACCGAGAUCUAC CAAGCCGGCUCCACCCCUUGCAACGGCGUGGAGGGCUUCA ACUGCUACUUCCCUCUGCAGAGCUACGGCUUCCAGCCCAC CAACGGCGUGGGCUACCAGCCCUACCGGGUGGUGGUGCU GAGCUUCGAGCUGCUGCACGCCCCAGCCACCGUGUGUGGC CCCAAGAAGAGCACCAACCUGGUGAAGAACAAGUGCGUG AACUUCAACUUCAACGGCCUUACCGGCACCGGCGUGCUG ACCGAGAGCAACAAGAAAUUCCUGCCCUUUCAGCAGUUC GGCCGGGACAUCGCCGACACCACCGACGCUGUGCGGGAUC CCCAGACCCUGGAGAUCCUGGACAUCACCCCUUGCAGCUU CGGCGGCGUGAGCGUGAUCACCCCAGGCACCAACACCAGC AACCAGGUGGCCGUGCUGUACCAGGACGUGAACUGCACC GAGGUGCCCGUGGCCAUCCACGCCGACCAGCUGACACCCA CCUGGCGGGUCUACAGCACCGGCAGCAACGUGUUCCAGA CCCGGGCCGGUUGCCUGAUCGGCGCCGAGCACGUGAACA ACAGCUACGAGUGCGACAUCCCCAUCGGCGCCGGCAUCUG UGCCAGCUACCAGACCCAGACCAAUUCACCCCGGAGGGCA AGGAGCGUGGCCAGCCAGAGCAUCAUCGCCUACACCAUG AGCCUGGGCGCCGAGAACAGCGUGGCCUACAGCAACAAC AGCAUCGCCAUCCCCACCAACUUCACCAUCAGCGUGACCA CCGAGAUUCUGCCCGUGAGCAUGACCAAGACCAGCGUGG ACUGCACCAUGUACAUCUGCGGCGACAGCACCGAGUGCA GCAACCUGCUGCUGCAGUACGGCAGCUUCUGCACCCAGCU GAACCGGGCCCUGACCGGCAUCGCCGUGGAGCAGGACAA GAACACCCAGGAGGUGUUCGCCCAGGUGAAGCAGAUCUA CAAGACCCCUCCCAUCAAGGACUUCGGCGGCUUCAACUUC AGCCAGAUCCUGCCCGACCCCAGCAAGCCCAGCAAGCGGA GCUUCAUCGAGGACCUGCUGUUCAACAAGGUGACCCUAG CCGACGCCGGCUUCAUCAAGCAGUACGGCGACUGCCUCGG CGACAUAGCCGCCCGGGACCUGAUCUGCGCCCAGAAGUUC	

	AACGGCCUGACCGUGCUGCCUCCCCUGCUGACCGACGAGA UGAUCGCCCAGUACACCAGCGCCCUGUUAGCCGGAACCAU CACCAGCGGCUGGACUUUCGGCGCUGGAGCCGCUCUGCA GAUCCCCUUCGCCAUGCAGAUGGCCUACCGGUUCAACGGC AUCGGCGUGACCCAGAACGUGCUGUACGAGAACCAGAAG CUGAUCGCCAACCAGUUCAACAGCGCCAUCGGCAAGAUCC AGGACAGCCUGAGCAGCACCGCUAGCGCCCUGGGCAAGC UGCAGGACGUGGUGAACCAGAACGCCCAGGCCCUGAACA CCCUGGUGAAGCAGCUGAGCAGCAACUUCGGCGCCAUCA GCAGCGUGCUGAACGACAUCCUGAGCCGGCUGGACCCUCC CGAGGCCGAGGUGCAGAUCGACCGGCUGAUCACUGGCCG GCUGCAGAGCCUGCAGACCUACGUGACCCAGCAGCUGAU CCGGGCCGCCGAGAUUCGGGCCAGCGCCAACCUGGCCGCC ACCAAGAUGAGCGAGUGCGUGCUGGGCCAGAGCAAGCGG GUGGACUUCUGCGGCAAGGGCUACCACCUGAUGAGCUUU CCCCAGAGCGCACCCCACGGAGUGGUGUUCCUGCACGUGA CCUACGUGCCCGCCCAGGAGAAGAACUUCACCACCGCCCC AGCCAUCUGCCACGACGGCAAGGCCCACUUUCCCCGGGAG GGCGUGUUCGUGAGCAACGGCACCCACUGGUUCGUGACC CAGCGGAACUUCUACGAGCCCCAGAUCAUCACCACCGACA ACACCUUCGUGAGCGGCAACUGCGACGUGGUGAUCGGCA UCGUGAACAACACCGUGUACGAUCCCCUGCAGCCCGAGCU GGACAGCUUCAAGGAGGAGCUGGACAAGUACUUCAAGAA UCACACCAGCCCCGACGUGGACCUGGGCGACAUCAGCGGC AUCAACGCCAGCGUGGUGAACAUCCAGAAGGAGAUCGAU CGGCUGAACGAGGUGGCCAAGAACCUGAACGAGAGCCUG AUCGACCUGCAGGAGCUGGGCAAGUACGAGCAGUACAUC AAGUGGCCCUGGUACAUCUGGCUGGGCUUCAUCGCCGGC CUGAUCGCCAUCGUGAUGGUGACCAUCAUGCUGUGCUGC AUGACCAGCUGCUGCAGCUGCCUGAAGGGCUGUUGCAGC UGCGGCAGCUGCUGCAAGUUCGACGAGGACGACAGCGAG CCCGUGCUGAAGGGCGUGAAGCUGCACUACACC	
3′ UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCC CCUUGGGCCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACC CGUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGG C	4
Corresponding amino acid sequence	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVF RSSVLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPF NDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKV CEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVS QPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRD LPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWT AGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKC TLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATR FASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLN DLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTG CVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQ AGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFE LLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKK FLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTN TSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQT RAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPRRARSVA SQSIIAYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTK TSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDK NTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDL LFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPL LTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRF	29

	NGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQ DVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDPPEAEVQ IDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLG QSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTT APAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNT FVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPD VDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYE QYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCLKGCCSC GSCCKFDEDDSEPVLKGVKLHYT	
PolyA tail	100 nt	
SARS-CoV-2 S Protein Variant 10		
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	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUA AGACCCCGGCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGUGAGCAGC CAGUGCGUGAACCUGACCACCCGGACCCAGCUGCCACCAG CCUACACCAACAGCUUCACCCGGGGCGUCUACUACCCCGA CAAGGUGUUCCGGAGCAGCGUCCUGCACAGCACCCAGGA CCUGUUCCUGCCCUUCUUCAGCAACGUGACCUGGUUCCAC GCCAUCCACGUGAGCGGCACCAACGGCACCAAGCGGUUCG ACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUCGC CAGCACCGAGAAGAGCAACAUCAUCCGGGGCUGGAUCUU CGGCACCACCCUGGACAGCAAGACCCAGAGCCUGCUGAUC GUGAAUAACGCCACCAACGUGGUGAUCAAGGUGUGCGAG UUCCAGUUCUGCAACGACCCCUUCCUGGGCGUGUACUACC ACAAGAACAACAAGAGCUGGAUGGAGAGCGAGUUCCGGG UGUACAGCAGCGCCAACAACUGCACCUUCGAGUACGUGA GCCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAGGGCA ACUUCAAGAACCUGCGGGAGUUCGUGUUCAAGAACAUCG ACGGCUACUUCAAGAUCUACAGCAAGCACACCCCAAUCA ACCUGGUGCGGGAUCUGCCCCAGGGCUUCUCAGCCCUGG AGCCCCUGGUGGACCUGCCCAUCGGCAUCAACAUCACCCG GUUCCAGACCCUGCUGGCCCUGCACCGGAGCUACCUGACC CCAGGCGACAGCAGCAGCGGGUGGACAGCAGGCGCGGCU GCUUACUACGUGGGCUACCUGCAGCCCCGGACCUUCCUGC UGAAGUACAACGAGAACGGCACCAUCACCGACGCCGUGG ACUGCGCCCUGGACCCUCUGAGCGAGACCAAGUGCACCCU GAAGAGCUUCACCGUGGAGAAGGGCAUCUACCAGACCAG CAACUUCCGGGUGCAGCCCACCGAGAGCAUCGUGCGGUU CCCCAACAUCACCAACCUGUGCCCCUUCGGCGAGGUGUUC AACGCCACCCGGUUCGCCAGCGUGUACGCCUGGAACCGGA AGCGGAUCAGCAACUGCGUGGCCGACUACAGCGUGCUGU ACAACAGCGCCAGCUUCAGCACCUUCAAGUGCUACGGCG UGAGCCCCACCAAGCUGAACGACCUGUGCUUCACCAACGU GUACGCCGACAGCUUCGUGAUCCGUGGCGACGAGGUGCG GCAGAUCGCACCCGGCCAGACAGGCAAGAUCGCCGACUAC AACUACAAGCUGCCCGACGACUUCACCGGCUGCGUGAUC GCCUGGAACAGCAACAACCUCGACAGCAAGGUGGGCGGC AACUACAACUACCUGUACCGGCUGUUCCGGAAGAGCAAC CUGAAGCCCUUCGAGCGGGACAUCAGCACCGAGAUCUAC	52

	CAAGCCGGCUCCACCCCUUGCAACGGCGUGGAGGGCUUCA ACUGCUACUUCCCUCUGCAGAGCUACGGCUUCCAGCCCAC CAACGGCGUGGGCUACCAGCCCUACCGGGUGGUGGUGCU GAGCUUCGAGCUGCUGCACGCCCCAGCCACCGUGUGUGGC CCCAAGAAGAGCACCAACCUGGUGAAGAACAAGUGCGUG AACUUCAACUUCAACGGCCUUACCGGCACCGGCGUGCUG ACCGAGAGCAACAAGAAAUUCCUGCCCUUUCAGCAGUUC GGCCGGGACAUCGCCGACACCACCGACGCUGUGCGGGAUC CCCAGACCCUGGAGAUCCUGGACAUCACCCCUUGCAGCUU CGGCGGCGUGAGCGUGAUCACCCCAGGCACCAACACCAGC AACCAGGUGGCCGUGCUGUACCAGGACGUGAACUGCACC GAGGUGCCCGUGGCCAUCCACGCCGACCAGCUGACACCCA CCUGGCGGGUCUACAGCACCGGCAGCAACGUGUUCCAGA CCCGGGCCGGUUGCCUGAUCGGCGCCGAGCACGUGAACA ACAGCUACGAGUGCGACAUCCCCAUCGGCGCCGGCAUCUG UGCCAGCUACCAGACCCAGACCAAUUCACCCCGGAGGGCA AGGAGCGUGGCCAGCCAGAGCAUCAUCGCCUACACCAUG AGCCUGGGCGCCGAGAACAGCGUGGCCUACAGCAACAAC AGCAUCGCCAUCCCCACCAACUUCACCAUCAGCGUGACCA CCGAGAUUCUGCCCGUGAGCAUGACCAAGACCAGCGUGG ACUGCACCAUGUACAUCUGCGGCGACAGCACCGAGUGCA GCAACCUGCUGCUGCAGUACGGCAGCUUCUGCACCCAGCU GAACCGGGCCCUGACCGGCAUCGCCGUGGAGCAGGACAA GAACACCCAGGAGGUGUUCGCCCAGGUGAAGCAGAUCUA CAAGACCCCUCCCAUCAAGGACUUCGGCGGCUUCAACUUC AGCCAGAUCCUGCCCGACCCCAGCAAGCCCAGCAAGCGGA GCUUCAUCGAGGACCUGCUGUUCAACAAGGUGACCCUAG CCGACGCCGGCUUCAUCAAGCAGUACGGCGACUGCCUCGG CGACAUAGCCGCCCGGGACCUGAUCUGCGCCCAGAAGUUC AACGGCCUGACCGUGCUGCCUCCCCUGCUGACCGACGAGA UGAUCGCCCAGUACACCAGCGCCCUGUUAGCCGGAACCAU CACCAGCGGCUGGACUUUCGGCGCUGGAGCCGCUCUGCA GAUCCCCUUCGCCAUGCAGAUGGCCUACCGGUUCAACGGC AUCGGCGUGACCCAGAACGUGCUGUACGAGAACCAGAAG CUGAUCGCCAACCAGUUCAACAGCGCCAUCGGCAAGAUCC AGGACAGCCUGAGCAGCACCGCUAGCGCCCUGGGCAAGC UGCAGGACGUGGUGAACCAGAACGCCCAGGCCCUGAACA CCCUGGUGAAGCAGCUGAGCAGCAACUUCGGCGCCAUCA GCAGCGUGCUGAACGACAUCCUGAGCCGGCUGGACAAGG UGGAGGCCGAGGUGCAGAUCGACCGGCUGAUCACUGGCC GGCUGCAGAGCCUGCAGACCUACGUGACCCAGCAGCUGA UCCGGGCCGCCGAGAUUCGGGCCAGCGCCAACCUGGCCGC CACCAAGAUGAGCGAGUGCGUGCUGGGCCAGAGCAAGCG GGUGGACUUCUGCGGCAAGGGCUACCACCUGAUGAGCUU UCCCCAGAGCGCACCCCACGGAGUGGUGUUCCUGCACGUG ACCUACGUGCCCGCCCAGGAGAAGAACUUCACCACCGCCC CAGCCAUCUGCCACGACGGCAAGGCCCACUUUCCCCGGGA GGGCGUGUUCGUGAGCAACGGCACCCACUGGUUCGUGAC CCAGCGGAACUUCUACGAGCCCCAGAUCAUCACCACCGAC AACACCUUCGUGAGCGGCAACUGCGACGUGGUGAUCGGC AUCGUGAACAACACCGUGUACGAUCCCCUGCAGCCCGAGC UGGACAGCUUCAAGGAGGAGCUGGACAAGUACUUCAAGA AUCACACCAGCCCCGACGUGGACCUGGGCGACAUCAGCGG CAUCAACGCCAGCGUGGUGAACAUCCAGAAGGAGAUCGA UCGGCUGAACGAGGUGGCCAAGAACCUGAACGAGAGCCU GAUCGACCUGCAGGAGCUGGGCAAGUACGAGCAGUACAU CAAGUGGCCCUGGUACAUCUGGCUGGGCUUCAUCGCCGG CCUGAUCGCCAUCGUGAUGGUGACCAUCAUGCUGUGCUG	

	CAUGACCAGCUGCUGCAGCUGCCUGAAGGGCUGUUGCAG CUGCGGCAGCUGCUGCAAGUUCGACGAGGACGAC	
3′ UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCC CCUUGGGCCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACC CGUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGG C	4
Corresponding amino acid sequence	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVF RSSVLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPF NDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKV CEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVS QPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRD LPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWT AGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKC TLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATR FASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLN DLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTG CVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQ AGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFE LLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKK FLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTN TSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQT RAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPRRARSVA SQSIIAYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTK TSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDK NTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDL LFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPL LTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRF NGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQ DVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEV QIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVL GQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNF TTAPAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTD NTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTS PDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGK YEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCLKGCC SCGSCCKFDEDD	33
PolyA tail	100 nt	
SARS-CoV-2 S Protein Variant 11		
SEQ ID NO: 53 consists of from 5′ end to 3′ end: 5′ UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 54, and 3′ UTR SEQ ID NO: 4.		53
Chemistry	1-methylpseudouridine	
Cap	7mG(5′)ppp(5′)NlmpNp	
5′ UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUA AGACCCCGGCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGUGAGCAGC CAGUGCGUGAACCUGACCACCCGGACCCAGCUGCCACCAG CCUACACCAACAGCUUCACCCGGGGCGUCUACUACCCCGA CAAGGUGUUCCGGAGCAGCGUCCUGCACAGCACCCAGGA CCUGUUCCUGCCCUUCUUCAGCAACGUGACCUGGUUCCAC GCCAUCCACGUGAGCGGCACCAACGGCACCAAGCGGUUCG ACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUCGC CAGCACCGAGAAGAGCAACAUCAUCCGGGGCUGGAUCUU	54

	CGGCACCACCCUGGACAGCAAGACCCAGAGCCUGCUGAUC GUGAAUAACGCCACCAACGUGGUGAUCAAGGUGUGCGAG UUCCAGUUCUGCAACGACCCCUUCCUGGGCGUGUACUACC ACAAGAACAACAAGAGCUGGAUGGAGAGCGAGUUCCGGG UGUACAGCAGCGCCAACAACUGCACCUUCGAGUACGUGA GCCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAGGGCA ACUUCAAGAACCUGCGGGAGUUCGUGUUCAAGAACAUCG ACGGCUACUUCAAGAUCUACAGCAAGCACACCCCAAUCA ACCUGGUGCGGGAUCUGCCCCAGGGCUUCUCAGCCCUGG AGCCCCUGGUGGACCUGCCCAUCGGCAUCAACAUCACCCG GUUCCAGACCCUGCUGGCCCUGCACCGGAGCUACCUGACC CCAGGCGACAGCAGCAGCGGGUGGACAGCAGGCGCGGCU GCUUACUACGUGGGCUACCUGCAGCCCCGGACCUUCCUGC UGAAGUACAACGAGAACGGCACCAUCACCGACGCCGUGG ACUGCGCCCUGGACCCUCUGAGCGAGACCAAGUGCACCCU GAAGAGCUUCACCGUGGAGAAGGGCAUCUACCAGACCAG CAACUUCCGGGUGCAGCCCACCGAGAGCAUCGUGCGGUU CCCCAACAUCACCAACCUGUGCCCCUUCGGCGAGGUGUUC AACGCCACCCGGUUCGCCAGCGUGUACGCCUGGAACCGGA AGCGGAUCAGCAACUGCGUGGCCGACUACAGCGUGCUGU ACAACAGCGCCAGCUUCAGCACCUUCAAGUGCUACGGCG UGAGCCCCACCAAGCUGAACGACCUGUGCUUCACCAACGU GUACGCCGACAGCUUCGUGAUCCGUGGCGACGAGGUGCG GCAGAUCGCACCCGGCCAGACAGGCAAGAUCGCCGACUAC AACUACAAGCUGCCCGACGACUUCACCGGCUGCGUGAUC GCCUGGAACAGCAACAACCUCGACAGCAAGGUGGGCGGC AACUACAACUACCUGUACCGGCUGUUCCGGAAGAGCAAC CUGAAGCCCUUCGAGCGGGACAUCAGCACCGAGAUCUAC CAAGCCGGCUCCACCCCUUGCAACGGCGUGGAGGGCUUCA ACUGCUACUUCCCUCUGCAGAGCUACGGCUUCCAGCCCAC CAACGGCGUGGGCUACCAGCCCUACCGGGUGGUGGUGCU GAGCUUCGAGCUGCUGCACGCCCCAGCCACCGUGUGUGGC CCCAAGAAGAGCACCAACCUGGUGAAGAACAAGUGCGUG AACUUCAACUUCAACGGCCUUACCGGCACCGGCGUGCUG ACCGAGAGCAACAAGAAAUUCCUGCCCUUUCAGCAGUUC GGCCGGGACAUCGCCGACACCACCGACGCUGUGCGGGAUC CCCAGACCCUGGAGAUCCUGGACAUCACCCCUUGCAGCUU CGGCGGCGUGAGCGUGAUCACCCCAGGCACCAACACCAGC AACCAGGUGGCCGUGCUGUACCAGGACGUGAACUGCACC GAGGUGCCCGUGGCCAUCCACGCCGACCAGCUGACACCCA CCUGGCGGGUCUACAGCACCGGCAGCAACGUGUUCCAGA CCCGGGCCGGUUGCCUGAUCGGCGCCGAGCACGUGAACA ACAGCUACGAGUGCGACAUCCCCAUCGGCGCCGGCAUCUG UGCCAGCUACCAGACCCAGACCAAUUCACCCCGGAGGGCA AGGAGCGUGGCCAGCCAGAGCAUCAUCGCCUACACCAUG AGCCUGGGCGCCGAGAACAGCGUGGCCUACAGCAACAAC AGCAUCGCCAUCCCCACCAACUUCACCAUCAGCGUGACCA CCGAGAUUCUGCCCGUGAGCAUGACCAAGACCAGCGUGG ACUGCACCAUGUACAUCUGCGGCGACAGCACCGAGUGCA GCAACCUGCUGCUGCAGUACGGCAGCUUCUGCACCCAGCU GAACCGGGCCCUGACCGGCAUCGCCGUGGAGCAGGACAA GAACACCCAGGAGGUGUUCGCCCAGGUGAAGCAGAUCUA CAAGACCCCUCCCAUCAAGGACUUCGGCGGCUUCAACUUC AGCCAGAUCCUGCCCGACCCCAGCAAGCCCAGCAAGCGGA GCUUCAUCGAGGACCUGCUGUUCAACAAGGUGACCCUAG CCGACGCCGGCUUCAUCAAGCAGUACGGCGACUGCCUCGG CGACAUAGCCGCCCGGGACCUGAUCUGCGCCCAGAAGUUC AACGGCCUGACCGUGCUGCCUCCCCUGCUGACCGACGAGA	

	UGAUCGCCCAGUACACCAGCGCCCUGUUAGCCGGAACCAU CACCAGCGGCUGGACUUUCGGCGCUGGAGCCGCUCUGCA GAUCCCCUUCGCCAUGCAGAUGGCCUACCGGUUCAACGGC AUCGGCGUGACCCAGAACGUGCUGUACGAGAACCAGAAG CUGAUCGCCAACCAGUUCAACAGCGCCAUCGGCAAGAUCC AGGACAGCCUGAGCAGCACCGCUAGCGCCCUGGGCAAGC UGCAGGACGUGGUGAACCAGAACGCCCAGGCCCUGAACA CCCUGGUGAAGCAGCUGAGCAGCAACUUCGGCGCCAUCA GCAGCGUGCUGAACGACAUCCUGAGCCGGCUGGACCCUCC CGAGGCCGAGGUGCAGAUCGACCGGCUGAUCACUGGCCG GCUGCAGAGCCUGCAGACCUACGUGACCCAGCAGCUGAU CCGGGCCGCCGAGAUUCGGGCCAGCGCCAACCUGGCCGCC ACCAAGAUGAGCGAGUGCGUGCUGGGCCAGAGCAAGCGG GUGGACUUCUGCGGCAAGGGCUACCACCUGAUGAGCUUU CCCCAGAGCGCACCCCACGGAGUGGUGUUCCUGCACGUGA CCUACGUGCCCGCCCAGGAGAAGAACUUCACCACCGCCCC AGCCAUCUGCCACGACGGCAAGGCCCACUUUCCCCGGGAG GGCGUGUUCGUGAGCAACGGCACCCACUGGUUCGUGACC CAGCGGAACUUCUACGAGCCCCAGAUCAUCACCACCGACA ACACCUUCGUGAGCGGCAACUGCGACGUGGUGAUCGGCA UCGUGAACAACACCGUGUACGAUCCCCUGCAGCCCGAGCU GGACAGCUUCAAGGAGGAGCUGGACAAGUACUUCAAGAA UCACACCAGCCCCGACGUGGACCUGGGCGACAUCAGCGGC AUCAACGCCAGCGUGGUGAACAUCCAGAAGGAGAUCGAU CGGCUGAACGAGGUGGCCAAGAACCUGAACGAGAGCCUG AUCGACCUGCAGGAGCUGGGCAAGUACGAGCAGUACAUC AAGUGGCCCUGGUACAUCUGGCUGGGCUUCAUCGCCGGC CUGAUCGCCAUCGUGAUGGUGACCAUCAUGCUGUGCUGC AUGACCAGCUGCUGCAGCUGCCUGAAGGGCUGUUGCAGC UGCGGCAGCUGCUGCAAGUUCGACGAGGACGAC	
3′ UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCC CCUUGGGCCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACC CGUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGG C	4
Corresponding amino acid sequence	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVF RSSVLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPF NDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKV CEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVS QPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRD LPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWT AGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKC TLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATR FASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLN DLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTG CVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQ AGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFE LLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKK FLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTN TSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQT RAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPRRARSVA SQSIIAYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTK TSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDK NTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDL LFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPL LTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRF NGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQ DVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDPPEAEVQ	34

	IDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLG QSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTT APAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNT FVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPD VDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYE QYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCLKGCCSC GSCCKFDEDD	
PolyA tail	100 nt	
SARS-CoV-2 S Protein Variant 12		
SEQ ID NO: 55 consists of from 5′ end to 3′ end: 5′ UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 56, and 3′ UTR SEQ ID NO: 4.		55
Chemistry	1-methylpseudouridine	
Cap	7mG(5′)ppp(5′)NlmpNp	
5′ UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUA AGACCCCGGCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	UUUCAACGACGGCGUGUACUUCGCCAGCACCGAGAAGAG CAACAUCAUCCGGGGCUGGAUCUUCGGCACCACCCUGGAC AGCAAGACCCAGAGCCUGCUGAUCGUGAAUAACGCCACC AACGUGGUGAUCAAGGUGUGCGAGUUCCAGUUCUGCAAC GACCCCUUCCUGGGCGUGUACUACCACAAGAACAACAAG AGCUGGAUGGAGAGCGAGUUCCGGGUGUACAGCAGCGCC AACAACUGCACCUUCGAGUACGUGAGCCAGCCCUUCCUG AUGGACCUGGAGGGCAAGCAGGGCAACUUCAAGAACCUG CGGGAGUUCGUGUUCAAGAACAUCGACGGCUACUUCAAG AUCUACAGCAAGCACACCCCAAUCAACCUGGUGCGGGAU CUGCCCCAGGGCUUCUCAGCCCUGGAGCCCCUGGUGGACC UGCCCAUCGGCAUCAACAUCACCCGGUUCCAGACCCUGCU GGCCCUGCACCGGAGCUACCUGACCCCAGGCGACAGCAGC AGCGGGUGGACAGCAGGCGCGGCUGCUUACUACGUGGGC UACCUGCAGCCCCGGACCUUCCUGCUGAAGUACAACGAG AACGGCACCAUCACCGACGCCGUGGACUGCGCCCUGGACC CUCUGAGCGAGACCAAGUGCACCCUGAAGAGCUUCACCG UGGAGAAGGGCAUCUACCAGACCAGCAACUUCCGGGUGC AGCCCACCGAGAGCAUCGUGCGGUUCCCCAACAUCACCAA CCUGUGCCCCUUCGGCGAGGUGUUCAACGCCACCCGGUUC GCCAGCGUGUACGCCUGGAACCGGAAGCGGAUCAGCAAC UGCGUGGCCGACUACAGCGUGCUGUACAACAGCGCCAGC UUCAGCACCUUCAAGUGCUACGGCGUGAGCCCCACCAAGC UGAACGACCUGUGCUUCACCAACGUGUACGCCGACAGCU UCGUGAUCCGUGGCGACGAGGUGCGGCAGAUCGCACCCG GCCAGACAGGCAAGAUCGCCGACUACAACUACAAGCUGC CCGACGACUUCACCGGCUGCGUGAUCGCCUGGAACAGCA ACAACCUCGACAGCAAGGUGGGCGGCAACUACAACUACC UGUACCGGCUGUUCCGGAAGAGCAACCUGAAGCCCUUCG AGCGGGACAUCAGCACCGAGAUCUACCAAGCCGGCUCCAC CCCUUGCAACGGCGUGGAGGGCUUCAACUGCUACUUCCC UCUGCAGAGCUACGGCUUCCAGCCCACCAACGGCGUGGGC UACCAGCCCUACCGGGUGGUGGUGCUGAGCUUCGAGCUG CUGCACGCCCCAGCCACCGUGUGUGGCCCCAAGAAGAGCA CCAACCUGGUGAAGAACAAGUGCGUGAACUUCAACUUCA ACGGCCUUACCGGCACCGGCGUGCUGACCGAGAGCAACA AGAAAUUCCUGCCCUUUCAGCAGUUCGGCCGGGACAUCG CCGACACCACCGACGCUGUGCGGGAUCCCCAGACCCUGGA	56

	GAUCCUGGACAUCACCCCUUGCAGCUUCGGCGGCGUGAG CGUGAUCACCCCAGGCACCAACACCAGCAACCAGGUGGCC GUGCUGUACCAGGACGUGAACUGCACCGAGGUGCCCGUG GCCAUCCACGCCGACCAGCUGACACCCACCUGGCGGGUCU ACAGCACCGGCAGCAACGUGUUCCAGACCCGGGCCGGUU GCCUGAUCGGCGCCGAGCACGUGAACAACAGCUACGAGU GCGACAUCCCCAUCGGCGCCGGCAUCUGUGCCAGCUACCA GACCCAGACCAAUUCACCCGGCAGCGGCGGCAGCGUGGCC AGCCAGAGCAUCAUCGCCUACACCAUGAGCCUGGGCGCCG AGAACAGCGUGGCCUACAGCAACAACAGCAUCGCCAUCCC CACCAACUUCACCAUCAGCGUGACCACCGAGAUUCUGCCC GUGAGCAUGACCAAGACCAGCGUGGACUGCACCAUGUAC AUCUGCGGCGACAGCACCGAGUGCAGCAACCUGCUGCUG CAGUACGGCAGCUUCUGCACCCAGCUGAACCGGGCCCUGA CCGGCAUCGCCGUGGAGCAGGACAAGAACACCCAGGAGG UGUUCGCCCAGGUGAAGCAGAUCUACAAGACCCCUCCCA UCAAGGACUUCGGCGGCUUCAACUUCAGCCAGAUCCUGC CCGACCCCAGCAAGCCCAGCAAGCGGAGCUUCAUCGAGGA CCUGCUGUUCAACAAGGUGACCCUAGCCGACGCCGGCUUC AUCAAGCAGUACGGCGACUGCCUCGGCGACAUAGCCGCCC GGGACCUGAUCUGCGCCCAGAAGUUCAACGGCCUGACCG UGCUGCCUCCCCUGCUGACCGACGAGAUGAUCGCCCAGUA CACCAGCGCCCUGUUAGCCGGAACCAUCACCAGCGGCUGG ACUUUCGGCGCUGGAGCCGCUCUGCAGAUCCCCUUCGCCA UGCAGAUGGCCUACCGGUUCAACGGCAUCGGCGUGACCC AGAACGUGCUGUACGAGAACCAGAAGCUGAUCGCCAACC AGUUCAACAGCGCCAUCGGCAAGAUCCAGGACAGCCUGA GCAGCACCGCUAGCGCCCUGGGCAAGCUGCAGGACGUGG UGAACCAGAACGCCCAGGCCCUGAACACCCUGGUGAAGC AGCUGAGCAGCAACUUCGGCGCCAUCAGCAGCGUGCUGA ACGACAUCCUGAGCCGGCUGGACCCUCCCGAGGCCGAGGU GCAGAUCGACCGGCUGAUCACUGGCCGGCUGCAGAGCCU GCAGACCUACGUGACCCAGCAGCUGAUCCGGGCCGCCGAG AUUCGGGCCAGCGCCAACCUGGCCGCCACCAAGAUGAGCG AGUGCGUGCUGGGCCAGAGCAAGCGGGUGGACUUCUGCG GCAAGGGCUACCACCUGAUGAGCUUUCCCCAGAGCGCACC CCACGGAGUGGUGUUCCUGCACGUGACCUACGUGCCCGCC CAGGAGAAGAACUUCACCACCGCCCCAGCCAUCUGCCACG ACGGCAAGGCCCACUUUCCCCGGGAGGGCGUGUUCGUGA GCAACGGCACCCACUGGUUCGUGACCCAGCGGAACUUCU ACGAGCCCCAGAUCAUCACCACCGACAACACCUUCGUGAG CGGCAACUGCGACGUGGUGAUCGGCAUCGUGAACAACAC CGUGUACGAUCCCCUGCAGCCCGAGCUGGACAGCUUCAA GGAGGAGCUGGACAAGUACUUCAAGAAUCACACCAGCCC CGACGUGGACCUGGGCGACAUCAGCGGCAUCAACGCCAG CGUGGUGAACAUCCAGAAGGAGAUCGAUCGGCUGAACGA GGUGGCCAAGAACCUGAACGAGAGCCUGAUCGACCUGCA GGAGCUGGGCAAGUACGAGCAGUACAUCAAGUGGCCCUG GUACAUCUGGCUGGGCUUCAUCGCCGGCCUGAUCGCCAU CGUGAUGGUGACCAUCAUGCUGUGCUGCAUGACCAGCUG CUGCAGCUGCCUGAAGGGCUGUUGCAGCUGCGGCAGCUG CUGCAAGUUCGACGAGGACGAC	
3′ UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCC CCUUGGGCCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACC CGUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGG C	4

Corresponding amino acid sequence	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVF RSSVLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPF NDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKV CEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVS QPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRD LPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWT AGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKC TLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATR FASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLN DLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTG CVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQ AGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFE LLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKK FLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTN TSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQT RAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPGSGGSVA SQSIIAYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTK TSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDK NTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDL LFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPL LTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRF NGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQ DVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDPPEAEVQ IDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLG QSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTT APAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNT FVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPD VDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYE QYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCLKGCCSC GSCCKFDEDD	35
PolyA tail	100 nt	
WIV16 S Protein Variant 1		
SEQ ID NO: 57 consists of from 5′ end to 3′ end: 5′ UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 48, and 3′ UTR SEQ ID NO: 4.		57
Chemistry	1-methylpseudouridine	
Cap	7mG(5′)ppp(5′)NlmpNp	
5′ UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUA AGACCCCGGCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUUUAUCUUCCUGUUCUUCCUGACCCUGACCAGCGGC AGCGACCUGGAAAGCUGCACCACCUUCGACGACGUGCAG GCCCCCAACUACCCUCAGCACAGCUCUAGCAGACGGGGCG UGUACUACCCCGACGAGAUCUUCAGAAGCGACACCCUGU ACCUGACCCAGGACCUGUUCCUGCCCUUCUACAGCAACGU GACCGGCUUCCACACCAUCAACCACAGAUUCGACAACCCC GUGAUCCCCUUCAAGGACGGGGUGUACUUUGCCGCCACC GAGAAGUCCAAUGUCGUGCGGGGAUGGGUGUUCGGCAGC ACCAUGAACAACAAGAGCCAGAGCGUGAUCAUCAUCAAC AACAGCACCAACGUCGUGAUCCGGGCCUGCAACUUCGAG CUGUGCGACAACCCAUUCUUCGCCGUGUCCAAGCCCACCG GCACCCAGACCCACACCAUGAUCUUCGACAACGCCUUCAA CUGCACCUUCGAGUACAUCAGCGACAGCUUCAGCCUGGA CGUGGCCGAGAAAAGCGGCAACUUCAAGCACCUGAGAGA AUUCGUGUUCAAGAACAAGGACGGCUUCCUGUACGUGUA	48

	CAAGGGCUACCAGCCCAUCGACGUCGUGCGCGAUCUGCCC AGCGGCUUCAACAUCCUGAAGCCCAUCUUCAAGCUGCCCC UGGGCAUCAACAUCACCAACUUCCGGGCUAUCCUGACCGC CUUCCUGCCCGCCCAGGAUACCUGGGGAACAAGCGCCGCU GCCUACUUCGUGGGCUACCUGAAGCCUGCCACCUUCAUGC UGAAGUACGACGAGAACGGCACCAUCACCGACGCCGUGG ACUGCAGCCAGAAUCCUCUGGCCGAGCUGAAGUGCAGCG UGAAGUCCUUCGAGAUCGACAAGGGCAUCUACCAGACCA GCAACUUCAGAGUGGCCCCCAGCAAAGAAGUCGUGCGGU UCCCCAAUAUCACCAACCUGUGCCCCUUCGGCGAGGUGUU CAACGCCACCACCUUUCCCAGCGUGUACGCCUGGGAGCGG AAGCGGAUCAGCAACUGCGUGGCCGACUACAGCGUGCUG UACAACUCCACCAGCUUCUCCACCUUCAAGUGCUACGGCG UGUCCGCCACCAAGCUGAACGACCUGUGCUUCAGCAAUG UGUACGCCGACUCCUUCGUCGUGAAGGGCGACGAUGUGC GCCAGAUCGCCCCUGGACAGACAGGCGUGAUCGCCGAUU ACAACUACAAGCUGCCUGACGACUUCACCGGCUGCGUGC UGGCCUGGAACACCAGAAACAUCGACGCCACCCAGACAG GCAACUACAAUUACAAGUACAGAAGCCUGCGGCACGGCA AGCUGCGGCCCUUCGAGAGGGACAUCUCCAACGUGCCCU UCAGCCCCGACGGCAAGCCUUGUACCCCCCCUGCCUUUAA CUGCUACUGGCCCCUGAACGACUACGGCUUCUACAUCACA AACGGCAUCGGCUAUCAGCCCUACCGGGUGGUGGUGCUG UCCUUUGAGCUGCUGAAUGCCCCUGCCACCGUGUGCGGCC CUAAGCUGAGCACCGACCUGAUCAAGAACCAGUGCGUGA ACUUCAACUUCAACGGCCUGACCGGCACCGGCGUGCUGAC ACCUAGCAGCAAGAGAUUCCAGCCCUUCCAGCAGUUCGG CCGGGACGUGCUGGAUUUCACCGACAGCGUGCGGGACCC CAAGACCAGCGAGAUCCUGGACAUCAGCCCCUGCAGCUUC GGCGGAGUGUCCGUGAUCACCCCCGGCACCAAUACCAGCU CUGAGGUGGCCGUGCUGUAUCAGGACGUGAACUGCACCG AUGUGCCCGUGGCCAUCCACGCCGAUCAGCUGACCCCAUC UUGGCGGGUGUACUCCACCGGCAACAACGUGUUCCAGAC ACAAGCCGGCUGCCUGAUCGGAGCCGAGCACGUGGACAC CAGCUACGAGUGCGACAUCCCUAUCGGCGCUGGCAUCUG CGCCAGCUACCACACCGUGUCCAGCCUGAGAAGCACCAGC CAGAAAUCUAUCGUGGCCUACACCAUGAGCCUGGGCGCC GACAGCUCUAUCGCCUACUCCAACAACACAAUCGCCAUCC CCACCAAUUUCAGCAUCUCCAUCACCACCGAAGUGAUGCC CGUGUCCAUGGCCAAGACCUCCGUGGAUUGCAACAUGUA CAUCUGCGGCGACAGCACCGAGUGCGCCAACCUGCUGCUG CAGUACGGCAGCUUCUGCACCCAGCUGAACAGAGCCCUG AGCGGAAUCGCCGUGGAACAGGACAGAAACACCCGGGAA GUGUUCGCCCAAGUGAAGCAGAUGUAUAAGACCCCCACC CUGAAGGAUUUCGGCGGCUUUAACUUCAGCCAGAUCCUG CCCGACCCUCUGAAGCCUACCAAGCGGAGCUUCAUCGAGG ACCUGCUGUUCAACAAAGUGACCCUGGCCGACGCCGGCU UUAUGAAGCAGUAUGGCGAGUGCCUGGGCGACAUCAACG CCCGGGAUCUGAUCUGCGCCCAGAAGUUUAACGGACUGA CCGUGCUGCCCCCUCUGCUGACCGACGAUAUGAUCGCCGC CUACACAGCCGCCCUGGUGUCUGGCACAGCUACCGCCGGA UGGACAUUUGGAGCUGGCGCCGCUCUGCAGAUCCCCUUU GCCAUGCAGAUGGCCUACCGGUUCAAUGGCAUCGGCGUG ACCCAGAAUGUGCUGUACGAGAACCAGAAGCAGAUCGCC AACCAGUUCAACAAGGCCAUUAGCCAGAUUCAGGAAAGC CUGACCACCACCAGCACCGCCCUGGGCAAACUGCAGGACG UCGUGAACCAGAACGCCCAGGCCCUGAACACCCUCGUGAA GCAGCUGAGCAGCAAUUUCGGCGCCAUCAGCUCCGUGCU	

	GAACGAUAUCCUGAGCAGACUGGACAAGGUGGAAGCAGA GGUGCAGAUCGACCGGCUGAUCACCGGCAGACUGCAGAG CCUGCAGACCUACGUGACACAGCAGCUGAUUAGAGCCGC CGAGAUCAGGGCCAGCGCCAAUCUGGCCGCCACAAAGAU GAGCGAGUGUGUGCUGGGCCAGAGCAAGCGGGUGGACUU CUGCGGCAAGGGCUAUCACCUGAUGAGCUUCCCCCAGGCC GCUCCUCACGGCGUGGUGUUUCUGCACGUGACAUACGUG CCCAGCCAGGAACGGAACUUCACCACCGCCCCAGCCAUCU GCCACGAGGGCAAGGCCUACUUCCCCCGGGAAGGCGUGU UCGUGUUUAACGGCACCUCCUGGUUUAUCACCCAGCGGA AUUUCUUCAGUCCGCAGAUCAUCACCACAGACAACACCU UCGUGUCCGGCAGCUGCGACGUCGUGAUUGGCAUCAUUA ACAACACCGUGUACGACCCCCUGCAGCCCGAGCUGGACAG CUUCAAAGAGGAACUGGACAAGUACUUCAAGAACCACAC CUCCCCCGACGUGGACCUGGGCGAUAUCUCCGGCAUCAAU GCCAGCGUCGUGAAUAUCCAGAAAGAGAUCGAUCGCCUG AACGAGGUGGCCAAGAACCUGAAUGAGAGCCUGAUCGAC CUGCAGGAACUGGGGAAGUACGAGCAGUACAUCAAGUGG CCUUGGUACGUGUGGCUGGGCUUUAUCGCCGGCCUGAUC GCCAUCGUGAUGGUCACCAUCCUGCUGUGCUGCAUGACC AGCUGUUGCAGCUGUCUGAAGGGCGCCUGCAGCUGUGGC UCCUGCUGCAAGUUCGAUGAGGACGACAGCGAGCCUGUG CUGAAAGGCGUGAAGCUGCACUACACC	
3′ UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCC CCUUGGGCCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACC CGUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGG C	4
Corresponding amino acid sequence	MFIFLFFLTLTSGSDLESCTTFDDVQAPNYPQHSSSRRGVYYPD EIFRSDTLYLTQDLFLPFYSNVTGFHTINHRFDNPVIPFKDGVY FAATEKSNVVRGWVFGSTMNNKSQSVIIINNSTNVVIRACNFE LCDNPFFAVSKPTGTQTHTMIFDNAFNCTFEYISDSFSLDVAEK SGNFKHLREFVFKNKDGFLYVYKGYQPIDVVRDLPSGFNILKP IFKLPLGINITNFRAILTAFLPAQDTWGTSAAAYFVGYLKPATF MLKYDENGTITDAVDCSQNPLAELKCSVKSFEIDKGIYQTSNF RVAPSKEVVRFPNITNLCPFGEVFNATTFPSVYAWERKRISNC VADYSVLYNSTSFSTFKCYGVSATKLNDLCFSNVYADSFVVK GDDVRQIAPGQTGVIADYNYKLPDDFTGCVLAWNTRNIDATQ TGNYNYKYRSLRHGKLRPFERDISNVPFSPDGKPCTPPAFNCY WPLNDYGFYITNGIGYQPYRVVVLSFELLNAPATVCGPKLSTD LIKNQCVNFNFNGLTGTGVLTPSSKRFQPFQQFGRDVLDFTDS VRDPKTSEILDISPCSFGGVSVITPGTNTSSEVAVLYQDVNCTD VPVAIHADQLTPSWRVYSTGNNVFQTQAGCLIGAEHVDTSYE CDIPIGAGICASYHTVSSLRSTSQKSIVAYTMSLGADSSIAYSN NTIAIPTNFSISITTEVMPVSMAKTSVDCNMYICGDSTECANLL LQYGSFCTQLNRALSGIAVEQDRNTREVFAQVKQMYKTPTLK DFGGFNFSQILPDPLKPTKRSFIEDLLFNKVTLADAGFMKQYG ECLGDINARDLICAQKFNGLTVLPPLLTDDMIAAYTAALVSGT ATAGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKQ IANQFNKAISQIQESLTTTSTALGKLQDVVNQNAQALNTLVKQ LSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTYVT QQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMS FPQAAPHGVVFLHVTYVPSQERNFTTAPAICHEGKAYFPREGV FVFNGTSWFITQRNFFSPQIITTDNTFVSGSCDVVIGIINNTVYD PLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEI DRLNEVAKNLNESLIDLQELGKYEQYIKWPWYVWLGFIAGLI AIVMVTILLCCMTSCCSCLKGACSCGSCCKFDEDDSEPVLKGV	47

	KLHYT	
PolyA tail	100 nt	
WIV16 S Protein Variant 2		
SEQ ID NO: 58 consists of from 5′ end to 3′ end: 5′ UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 50, and 3′ UTR SEQ ID NO: 4.		58
Chemistry	1-methylpseudouridine	
Cap	7mG(5′)ppp(5′)NlmpNp	
5′ UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUA AGACCCCGGCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUUCAUCUUCCUGUUCUUCCUGACCCUGACCAGCGGC AGCGACCUGGAGAGCUGCACCACCUUCGACGACGUGCAG GCCCCUAACUACCCUCAGCACAGCAGCAGCAGAAGAGGCG UGUACUACCCUGACGAGAUCUUCAGAAGCGACACCCUGU ACCUGACCCAGGACCUGUUCCUGCCUUUCUACAGCAACGU GACCGGCUUCCACACCAUCAACCACAGAUUCGACAACCCU GUGAUCCCUUUCAAGGACGGCGUGUACUUCGCCGCCACC GAGAAGAGCAACGUGGUGAGAGGCUGGGUGUUCGGCAGC ACCAUGAACAACAAGAGCCAGAGCGUGAUCAUCAUCAAC AACAGCACCAACGUGGUGAUCAGAGCCUGCAACUUCGAG CUGUGCGACAACCCUUUCUUCGCCGUGAGCAAGCCUACCG GCACCCAGACCCACACCAUGAUCUUCGACAACGCCUUCAA CUGCACCUUCGAGUACAUCAGCGACAGCUUCAGCCUGGA CGUGGCCGAGAAGAGCGGCAACUUCAAGCACCUGAGAGA GUUCGUGUUCAAGAACAAGGACGGCUUCCUGUACGUGUA CAAGGGCUACCAGCCUAUCGACGUGGUGAGAGACCUGCC UAGCGGCUUCAACAUCCUGAAGCCUAUCUUCAAGCUGCC UCUGGGCAUCAACAUCACCAACUUCAGAGCCAUCCUGACC GCCUUCCUGCCUGCCCAGGACACCUGGGGCACCAGCGCCG CCGCCUACUUCGUGGGCUACCUGAAGCCUGCCACCUUCAU GCUGAAGUACGACGAGAACGGCACCAUCACCGACGCCGU GGACUGCAGCCAGAACCCUCUGGCCGAGCUGAAGUGCAG CGUGAAGAGCUUCGAGAUCGACAAGGGCAUCUACCAGAC CAGCAACUUCAGAGUGGCCCCUAGCAAGGAGGUGGUGAG AUUCCCUAACAUCACCAACCUGUGCCCUUUCGGCGAGGU GUUCAACGCCACCACCUUCCCUAGCGUGUACGCCUGGGAG AGAAAGAGAAUCAGCAACUGCGUGGCCGACUACAGCGUG CUGUACAACAGCACCAGCUUCAGCACCUUCAAGUGCUAC GGCGUGAGCGCCACCAAGCUGAACGACCUGUGCUUCAGC AACGUGUACGCCGACAGCUUCGUGGUGAAGGGCGACGAC GUGAGACAGAUCGCCCCUGGCCAGACCGGCGUGAUCGCC GACUACAACUACAAGCUGCCUGACGACUUCACCGGCUGC GUGCUGGCCUGGAACACCAGAAACAUCGACGCCACCCAG ACCGGCAACUACAACUACAAGUACAGAAGCCUGAGACAC GGCAAGCUGAGACCUUUCGAGAGAGACAUCAGCAACGUG CCUUUCAGCCCUGACGGCAAGCCUUGCACCCCUCCUGCCU UCAACUGCUACUGGCCUCUGAACGACUACGGCUUCUACA UCACCAACGGCAUCGGCUACCAGCCUUACAGAGUGGUGG UGCUGAGCUUCGAGCUGCUGAACGCCCCUGCCACCGUGU GCGGCCCUAAGCUGAGCACCGACCUGAUCAAGAACCAGU GCGUGAACUUCAACUUCAACGGCCUGACCGGCACCGGCG UGCUGACCCCUAGCAGCAAGAGAUUCCAGCCUUUCCAGC AGUUCGGCAGAGACGUGCUGGACUUCACCGACAGCGUGA GAGACCCUAAGACCAGCGAGAUCCUGGACAUCAGCCCUU	50

	GCAGCUUCGGCGGCGUGAGCGUGAUCACCCCUGGCACCA ACACCAGCAGCGAGGUGGCCGUGCUGUACCAGGACGUGA ACUGCACCGACGUGCCUGUGGCCAUCCACGCCGACCAGCU GACCCCUAGCUGGAGAGUGUACAGCACCGGCAACAACGU GUUCCAGACCCAGGCCGGCUGCCUGAUCGGCGCCGAGCAC GUGGACACCAGCUACGAGUGCGACAUCCCUAUCGGCGCC GGCAUCUGCGCCAGCUACCACACCGUGAGCAGCCUGAGA AGCACCAGCCAGAAGAGCAUCGUGGCCUACACCAUGAGC CUGGGCGCCGACAGCAGCAUCGCCUACAGCAACAACACCA UCGCCAUCCCUACCAACUUCAGCAUCAGCAUCACCACCGA GGUGAUGCCUGUGAGCAUGGCCAAGACCAGCGUGGACUG CAACAUGUACAUCUGCGGCGACAGCACCGAGUGCGCCAA CCUGCUGCUGCAGUACGGCAGCUUCUGCACCCAGCUGAAC AGAGCCCUGAGCGGCAUCGCCGUGGAGCAGGACAGAAAC ACCAGAGAGGUGUUCGCCCAGGUGAAGCAGAUGUACAAG ACCCCUACCCUGAAGGACUUCGGCGGCUUCAACUUCAGCC AGAUCCUGCCUGACCCUCUGAAGCCUACCAAGAGAAGCU UCAUCGAGGACCUGCUGUUCAACAAGGUGACCCUGGCCG ACGCCGGCUUCAUGAAGCAGUACGGCGAGUGCCUGGGCG ACAUCAACGCCAGAGACCUGAUCUGCGCCCAGAAGUUCA ACGGCCUGACCGUGCUGCCUCCUCUGCUGACCGACGACAU GAUCGCCGCCUACACCGCCGCCCUGGUGAGCGGCACCGCC ACCGCCGGCUGGACCUUCGGCGCCGGCGCCGCCCUGCAGA UCCCUUUCGCCAUGCAGAUGGCCUACAGAUUCAACGGCA UCGGCGUGACCCAGAACGUGCUGUACGAGAACCAGAAGC AGAUCGCCAACCAGUUCAACAAGGCCAUCAGCCAGAUCC AGGAGAGCCUGACCACCACCAGCACCGCCCUGGGCAAGCU GCAGGACGUGGUGAACCAGAACGCCCAGGCCCUGAACAC CCUGGUGAAGCAGCUGAGCAGCAACUUCGGCGCCAUCAG CAGCGUGCUGAACGACAUCCUGAGCAGACUGGACCCUCC UGAGGCCGAGGUGCAGAUCGACAGACUGAUCACCGGCAG ACUGCAGAGCCUGCAGACCUACGUGACCCAGCAGCUGAU CAGAGCCGCCGAGAUCAGAGCCAGCGCCAACCUGGCCGCC ACCAAGAUGAGCGAGUGCGUGCUGGGCCAGAGCAAGAGA GUGGACUUCUGCGGCAAGGGCUACCACCUGAUGAGCUUC CCUCAGGCCGCCCCUCACGGCGUGGUGUUCCUGCACGUGA CCUACGUGCCUAGCCAGGAGAGAAACUUCACCACCGCCCC UGCCAUCUGCCACGAGGGCAAGGCCUACUUCCCUAGAGA GGGCGUGUUCGUGUUCAACGGCACCAGCUGGUUCAUCAC CCAGAGAAACUUCUUCAGCCCUCAGAUCAUCACCACCGAC AACACCUUCGUGAGCGGCAGCUGCGACGUGGUGAUCGGC AUCAUCAACAACACCGUGUACGACCCUCUGCAGCCUGAGC UGGACAGCUUCAAGGAGGAGCUGGACAAGUACUUCAAGA ACCACACCAGCCCUGACGUGGACCUGGGCGACAUCAGCGG CAUCAACGCCAGCGUGGUGAACAUCCAGAAGGAGAUCGA CAGACUGAACGAGGUGGCCAAGAACCUGAACGAGAGCCU GAUCGACCUGCAGGAGCUGGGCAAGUACGAGCAGUACAU CAAGUGGCCUUGGUACGUGUGGCUGGGCUUCAUCGCCGG CCUGAUCGCCAUCGUGAUGGUGACCAUCCUGCUGUGCUG CAUGACCAGCUGCUGCAGCUGCCUGAAGGGCGCCUGCAG CUGCGGCAGCUGCUGCAAGUUCGACGAGGACGACAGCGA GCCUGUGCUGAAGGGCGUGAAGCUGCACUACACC	
3′ UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCC CCUUGGGCCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACC CGUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGG C	4

Corresponding amino acid sequence	MFIFLFFLTLTSGSDLESCTTFDDVQAPNYPQHSSSRRGVYYPD EIFRSDTLYLTQDLFLPFYSNVTGFHTINHRFDNPVIPFKDGVY FAATEKSNVVRGWVFGSTMNNKSQSVIIINNSTNVVIRACNFE LCDNPFFAVSKPTGTQTHTMIFDNAFNCTFEYISDSFSLDVAEK SGNFKHLREFVFKNKDGFLYVYKGYQPIDVVRDLPSGFNILKP IFKLPLGINITNFRAILTAFLPAQDTWGTSAAAYFVGYLKPATF MLKYDENGTITDAVDCSQNPLAELKCSVKSFEIDKGIYQTSNF RVAPSKEVVRFPNITNLCPFGEVFNATTFPSVYAWERKRISNC VADYSVLYNSTSFSTFKCYGVSATKLNDLCFSNVYADSFVVK GDDVRQIAPGQTGVIADYNYKLPDDFTGCVLAWNTRNIDATQ TGNYNYKYRSLRHGKLRPFERDISNVPFSPDGKPCTPPAFNCY WPLNDYGFYITNGIGYQPYRVVVLSFELLNAPATVCGPKLSTD LIKNQCVNFNFNGLTGTGVLTPSSKRFQPFQQFGRDVLDFTDS VRDPKTSEILDISPCSFGGVSVITPGTNTSSEVAVLYQDVNCTD VPVAIHADQLTPSWRVYSTGNNVFQTQAGCLIGAEHVDTSYE CDIPIGAGICASYHTVSSLRSTSQKSIVAYTMSLGADSSIAYSN NTIAIPTNFSISITTEVMPVSMAKTSVDCNMYICGDSTECANLL LQYGSFCTQLNRALSGIAVEQDRNTREVFAQVKQMYKTPTLK DFGGFNFSQILPDPLKPTKRSFIEDLLFNKVTLADAGFMKQYG ECLGDINARDLICAQKFNGLTVLPPLLTDDMIAAYTAALVSGT ATAGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKQ IANQFNKAISQIQESLTTTSTALGKLQDVVNQNAQALNTLVKQ LSSNFGAISSVLNDILSRLDPPEAEVQIDRLITGRLQSLQTYVTQ QLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSF PQAAPHGVVFLHVTYVPSQERNFTTAPAICHEGKAYFPREGVF VFNGTSWFITQRNFFSPQIITTDNTFVSGSCDVVIGIINNTVYDP LQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEID RLNEVAKNLNESLIDLQELGKYEQYIKWPWYVWLGFIAGLIAI VMVTILLCCMTSCCSCLKGACSCGSCCKFDEDDSEPVLKGVK LHYT	49
PolyA tail	100 nt	
MERS S Protein Variant 1		
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	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUA AGACCCCGGCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGAUCCACAGCGUGUUCCUGCUGAUGUUCCUCCUUACC CCUACCGAGAGCUACGUGGACGUCGGCCCUGACAGCGUU AAGAGCGCUUGCAUCGAGGUGGACAUCCAGCAGACCUUC UUCGACAAGACCUGGCCUAGACCUAUCGACGUGAGCAAG GCCGACGGCAUCAUCUACCCUCAGGGCAGAACCUACAGCA ACAUCACCAUCACCUACCAGGGCCUGUUCCCUUAUCAGGG CGACCACGGCGACAUGUACGUGUACAGCGCCGGCCACGCC ACCGGCACAACGCCCCAGAAGCUUUUCGUGGCCAACUACU CCCAGGACGUGAAGCAGUUCGCCAACGGCUUCGUGGUGA GAAUCGGCGCCGCCGCCAACUCCACUGGAACCGUGAUCAU CAGCCCUAGCACCAGCGCCACCAUCAGAAAGAUUUAUCCU GCCUUUAUGCUGGGCUCCAGCGUGGGUAACUUCAGCGAC GGCAAGAUGGGCAGAUUCUUCAACCACACCCUGGUGCUG CUGCCUGACGGCUGCGGCACCCUGCUGAGAGCCUUCUACU GCAUCCUGGAGCCUAGAAGCGGCAACCACUGCCCUGCCGG	61

	CAACAGCUACACCAGCUUCGCAACCUAUCACACCCCUGCC ACCGACUGUUCUGACGGUAACUACAACAGAAACGCCAGC CUGAACAGCUUCAAGGAGUACUUCAACCUGAGAAACUGC ACCUUCAUGUACACCUAUAAUAUCACCGAGGACGAGAUC CUCGAGUGGUUCGGCAUAACCCAGACCGCCCAAGGCGUG CACCUGUUCAGCAGCAGAUACGUUGAUCUGUACGGCGGC AACAUGUUCCAGUUCGCUACCCUGCCUGUGUACGACACC AUCAAGUACUACAGCAUCAUCCCUCAUUCUAUUAGAAGC AUCCAGAGCGACAGAAAGGCCUGGGCCGCUUUCUACGUA UACAAGCUGCAGCCUCUCACAUUCUUGCUCGACUUCUCU GUGGACGGCUAUAUCCGCAGGGCCAUCGACUGCGGCUUC AACGACCUGAGCCAGCUGCACUGCAGCUACGAGAGCUUC GACGUGGAGAGCGGAGUUUAUUCCGUGAGCAGCUUCGAG GCCAAGCCUAGCGGCUCUGUAGUGGAGCAGGCCGAGGGC GUGGAGUGCGAUUUCAGCCCUCUGCUGAGCGGUACCCCU CCUCAGGUGUACAACUUCAAGAGACUGGUGUUCACGAAC UGCAACUACAAUCUGACCAAACUGCUUUCGCUUUUCUCC GUGAACGACUUCACCUGCAGCCAGAUUUCUCCGGCAGCC AUCGCCAGCAACUGCUACAGCAGCUUGAUCCUUGACUAC UUCAGCUACCCUCUGAGCAUGAAGUCCGACUUAAGUGUA UCCUCAGCCGGCCCUAUCAGCCAGUUCAAUUACAAGCAG AGCUUCAGCAACCCUACCUGCCUAAUUUUGGCCACCGUGC CUCACAACCUGACUACAAUUACCAAGCCACUCAAGUAUU CCUACAUUAACAAGUGUAGCCGAUUCCUGAGCGACGACA GAACCGAGGUGCCUCAGCUGGUGAACGCCAACCAGUACA GCCCUUGCGUGUCGAUCGUGCCAAGUACCGUGUGGGAGG ACGGCGACUACUACAGAAAGCAGCUGUCUCCUCUCGAAG GCGGCGGGUGGCUGGUGGCAAGCGGAAGCACAGUGGCCA UGACCGAGCAGCUGCAGAUGGGCUUCGGAAUUACCGUGC AGUACGGCACCGACACCAAUAGUGUCUGCCCUAAGCUGG AAUUCGCGAACGACACUAAGAUUGCCUCCCAACUGGGAA AUUGCGUAGAGUACUCUCUGUACGGAGUGUCCGGCAGAG GUGUCUUCCAGAAUUGCACAGCCGUGGGCGUGAGACAGC AGAGAUUCGUCUACGACGCCUACCAGAACCUGGUGGGCU AUUAUAGUGACGACGGCAACUACUACUGCCUGCGGGCCU GCGUUAGUGUGCCUGUCUCCGUUAUCUACGACAAGGAGA CAAAGACUCACGCCACACUUUUCGGAUCUGUCGCCUGCG AGCACAUCAGUAGUACCAUGUCUCAGUAUAGCAGAAGCA CCAGGUCUAUGCUGAAGAGACGGGACUCAACCUACGGAC CACUUCAGACCCCUGUGGGCUGCGUGCUGGGCCUCGUAA AUAGCUCUCUGUUUGUGGAGGACUGUAAACUGCCACUGG GCCAGAGCCUGUGUGCUUUACCUGACACACCUAGUACAC UGACACCAGCGAGCGUGGGUAGUGUACCAGGCGAGAUGA GACUGGCCAGCAUCGCUUUCAAUCACCCUAUCCAGGUGG ACCAGCUCAAUUCCUCUUACUUCAAGCUGAGCAUCCCUAC CAAUUUCUCUUUCGGCGUGACCCAGGAGUACAUCCAGAC CACAAUACAGAAGGUGACCGUAGAUUGCAAGCAGUACGU GUGUAACGGAUUCCAGAAGUGCGAGCAAUUGCUCAGGGA GUACGGCCAGUUCUGUAGCAAGAUCAACCAGGCUCUGCA CGGGGCCAAUCUGCGACAGGACGACAGCGUAAGAAACCU GUUCGCCAGCGUAAAGUCUAGCCAGUCGAGUCCAAUCAU ACCAGGCUUCGGCGGAGAUUUCAAUCUCACCUUAUUGGA GCCAGUUUCCAUCUCUACGGGUUCGAGGAGCGCUAGGAG CGCAAUCGAGGACCUGCUGUUCGAUAAGGUCACCAUCGC CGACCCUGGCUACAUGCAGGGCUACGACGACUGCAUGCA GCAGGGCCCAGCCUCCGCCAGAGAUCUGAUCUGCGCCCAG UACGUCGCCGGCUACAAGGUGCUGCCUCCUCUGAUGGAC GUUAACAUGGAGGCCGCCUAUACUAGUAGUCUUCUGGGA	

	AGCAUUGCAGGCGUGGGCUGGACCGCCGGCCUGUCUAGC UUCGCGGCCAUACCUUUCGCCCAGAGCAUCUUCUACAGAC UGAACGGUGUGGGCAUCACACAACAGGUACUGUCUGAGA AUCAGAAGCUGAUCGCCAACAAGUUCAAUCAGGCACUUG GCGCCAUGCAGACCGGCUUCACCACCACCAACGAGGCCUU CCACAAGGUCCAGGACGCCGUGAACAACAACGCUCAGGCC UUGAGCAAGUUAGCGAGCGAACUUAGCAACACCUUCGGC GCCAUCAGUGCAAGCAUUGGAGACAUUAUCCAGAGGCUC GACCCUCCUGAGCAGGACGCUCAGAUCGAUCGGUUGAUC AACGGCAGACUGACCACUCUGAACGCCUUCGUUGCCCAAC AACUGGUGCGGUCUGAGAGCGCCGCUUUAUCCGCCCAGC UGGCCAAGGACAAGGUUAACGAGUGCGUGAAGGCACAGU CGAAGCGUUCAGGAUUCUGCGGCCAGGGCACCCACAUCG UGAGCUUCGUCGUGAACGCCCCUAACGGCCUGUACUUCA UGCACGUCGGAUAUUACCCUAGCAACCAUAUUGAAGUGG UGAGCGCGUACGGCCUCUGUGACGCAGCUAAUCCUACAA ACUGCAUCGCCCCUGUGAACGGUUACUUCAUCAAGACCA ACAACACCAGAAUCGUGGACGAGUGGUCAUACACGGGCA GUUCAUUCUACGCCCCUGAGCCGAUCACUAGCCUUAACAC CAAGUACGUGGCCCCACAAGUGACAUACCAGAACAUUAG CACAAACCUGCCUCCACCGCUGUUAGGUAACAGCACGGGC AUCGACUUCCAGGACGAAUUAGACGAGUUCUUCAAGAAC GUGUCCACCAGCAUCCCAAACUUCGGCAGCCUGACCCAGA UCAACACAACCUUACUCGACCUGACCUACGAGAUGCUGA GCCUCCAGCAGGUUGUCAAGGCCCUCAACGAAUCAUAUA UCGACUUGAAGGAGCUUGGCAAUUACACUUACUACAACA AGUGGCCUUGGUACAUCUGGCUCGGCUUCAUCGCCGGGC UGGUCGCCCUCGCCCUGUGCGUCUUCUUCAUCCUGUGCUG CACAGGUUGUGGAACCAACUGUAUGGGCAAGCUGAAGUG CAACCGUUGCUGUGAUAGAUACGAGGAGUACGAUCUGGA ACCACAUAAGGUGCACGUGCAC	
3′ UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCC CCUUGGGCCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACC CGUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGG C	4
Corresponding amino acid sequence	MIHSVFLLMFLLTPTESYVDVGPDSVKSACIEVDIQQTFFDKT WPRPIDVSKADGIIYPQGRTYSNITITYQGLFPYQGDHGDMYV YSAGHATGTTPQKLFVANYSQDVKQFANGFVVRIGAAANST GTVIISPSTSATIRKIYPAFMLGSSVGNFSDGKMGRFFNHTLVL LPDGCGTLLRAFYCILEPRSGNHCPAGNSYTSFATYHTPATDC SDGNYNRNASLNSFKEYFNLRNCTFMYTYNITEDEILEWFGIT QTAQGVHLFSSRYVDLYGGNMFQFATLPVYDTIKYYSIIPHSI RSIQSDRKAWAAFYVYKLQPLTFLLDFSVDGYIRRAIDCGFND LSQLHCSYESFDVESGVYSVSSFEAKPSGSVVEQAEGVECDFS PLLSGTPPQVYNFKRLVFTNCNYNLTKLLSLFSVNDFTCSQISP AAIASNCYSSLILDYFSYPLSMKSDLSVSSAGPISQFNYKQSFS NPTCLILATVPHNLTTITKPLKYSYINKCSRFLSDDRTEVPQLV NANQYSPCVSIVPSTVWEDGDYYRKQLSPLEGGGWLVASGST VAMTEQLQMGFGITVQYGTDTNSVCPKLEFANDTKIASQLGN CVEYSLYGVSGRGVFQNCTAVGVRQQRFVYDAYQNLVGYYS DDGNYYCLRACVSVPVSVIYDKETKTHATLFGSVACEHISST MSQYSRSTRSMLKRRDSTYGPLQTPVGCVLGLVNSSLFVEDC KLPLGQSLCALPDTPSTLTPASVGSVPGEMRLASIAFNHPIQVD QLNSSYFKLSIPTNFSFGVTQEYIQTTIQKVTVDCKQYVCNGF QKCEQLLREYGQFCSKINQALHGANLRQDDSVRNLFASVKSS QSSPIIPGFGGDFNLTLLEPVSISTGSRSARSAIEDLLFDKVTIAD PGYMQGYDDCMQQGPASARDLICAQYVAGYKVLPPLMDVN	59

	MEAAYTSSLLGSIAGVGWTAGLSSFAAIPFAQSIFYRLNGVGIT QQVLSENQKLIANKFNQALGAMQTGFTTTNEAFHKVQDAVN NNAQALSKLASELSNTFGAISASIGDIIQRLDPPEQDAQIDRLIN GRLTTLNAFVAQQLVRSESAALSAQLAKDKVNECVKAQSKR SGFCGQGTHIVSFVVNAPNGLYFMHVGYYPSNHIEVVSAYGL CDAANPTNCIAPVNGYFIKTNNTRIVDEWSYTGSSFYAPEPITS LNTKYVAPQVTYQNISTNLPPPLLGNSTGIDFQDELDEFFKNV STSIPNFGSLTQINTTLLDLTYEMLSLQQVVKALNESYIDLKEL GNYTYYNKWPWYIWLGFIAGLVALALCVFFILCCTGCGTNC MGKLKCNRCCDRYEEYDLEPHKVHVH	
PolyA tail	100 nt	
SARS-CoV-2 Variant 25		
SEQ ID NO: 86 consists of from 5′ end to 3′ end: 5′ UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 62, and 3′ UTR SEQ ID NO: 4.		86
Chemistry	1-methylpseudouridine	
Cap	7mG(5′)ppp(5′)NlmpNp	
5′ UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUA AGACCCCGGCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGUGAGCAGC CAGUGCGUGAACCUGACCACCCGGACCCAGCUGCCACCAG CCUACACCAACAGCUUCACCCGGGGCGUCUACUACCCCGA CAAGGUGUUCCGGAGCAGCGUCCUGCACAGCACCCAGGA CCUGUUCCUGCCCUUCUUCAGCAACGUGACCUGGUUCCAC GCCAUCCACGUGAGCGGCACCAACGGCACCAAGCGGUUCG ACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUCGC CAGCACCGAGAAGAGCAACAUCAUCCGGGGCUGGAUCUU CGGCACCACCCUGGACAGCAAGACCCAGAGCCUGCUGAUC GUGAAUAACGCCACCAACGUGGUGAUCAAGGUGUGCGAG UUCCAGUUCUGCAACGACCCCUUCCUGGGCGUGUACUACC ACAAGAACAACAAGAGCUGGAUGGAGAGCGAGUUCCGGG UGUACAGCAGCGCCAACAACUGCACCUUCGAGUACGUGA GCCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAGGGCA ACUUCAAGAACCUGCGGGAGUUCGUGUUCAAGAACAUCG ACGGCUACUUCAAGAUCUACAGCAAGCACACCCCAAUCA ACCUGGUGCGGGAUCUGCCCCAGGGCUUCUCAGCCCUGG AGCCCCUGGUGGACCUGCCCAUCGGCAUCAACAUCACCCG GUUCCAGACCCUGCUGGCCCUGCACCGGAGCUACCUGACC CCAGGCGACAGCAGCAGCGGGUGGACAGCAGGCGCGGCU GCUUACUACGUGGGCUACCUGCAGCCCCGGACCUUCCUGC UGAAGUACAACGAGAACGGCACCAUCACCGACGCCGUGG ACUGCGCCCUGGACCCUCUGAGCGAGACCAAGUGCACCCU GAAGAGCUUCACCGUGGAGAAGGGCAUCUACCAGACCAG CAACUUCCGGGUGCAGCCCACCGAGAGCAUCGUGCGGUU CCCCAACAUCACCAACCUGUGCCCCUUCGGCGAGGUGUUC AACGCCACCCGGUUCGCCAGCGUGUACGCCUGGAACCGGA AGCGGAUCAGCAACUGCGUGGCCGACUACAGCGUGCUGU ACAACAGCGCCAGCUUCAGCACCUUCAAGUGCUACGGCG UGAGCCCCACCAAGCUGAACGACCUGUGCUUCACCAACGU GUACGCCGACAGCUUCGUGAUCCGUGGCGACGAGGUGCG GCAGAUCGCACCCGGCCAGACAGGCAAGAUCGCCGACUAC AACUACAAGCUGCCCGACGACUUCACCGGCUGCGUGAUC GCCUGGAACAGCAACAACCUCGACAGCAAGGUGGGCGGC AACUACAACUACCUGUACCGGCUGUUCCGGAAGAGCAAC CUGAAGCCCUUCGAGCGGGACAUCAGCACCGAGAUCUAC CAAGCCGGCUCCACCCCUUGCAACGGCGUGGAGGGCUUCA	62

	ACUGCUACUUCCCUCUGCAGAGCUACGGCUUCCAGCCCAC CAACGGCGUGGGCUACCAGCCCUACCGGGUGGUGGUGCU GAGCUUCGAGCUGCUGCACGCCCCAGCCACCGUGUGUGGC CCCAAGAAGAGCACCAACCUGGUGAAGAACAAGUGCGUG AACUUCAACUUCAACGGCCUUACCGGCACCGGCGUGCUG ACCGAGAGCAACAAGAAAUUCCUGCCCUUUCAGCAGUUC GGCCGGGACAUCGCCGACACCACCGACGCUGUGCGGGAUC CCCAGACCCUGGAGAUCCUGGACAUCACCCCUUGCAGCUU CGGCGGCGUGAGCGUGAUCACCCCAGGCACCAACACCAGC AACCAGGUGGCCGUGCUGUACCAGGACGUGAACUGCACC GAGGUGCCCGUGGCCAUCCACGCCGACCAGCUGACACCCA CCUGGCGGGUCUACAGCACCGGCAGCAACGUGUUCCAGA CCCGGGCCGGUUGCCUGAUCGGCGCCGAGCACGUGAACA ACAGCUACGAGUGCGACAUCCCCAUCGGCGCCGGCAUCUG UGCCAGCUACCAGACCCAGACCGUGUCACUGAGGAGCGU GGCCAGCCAGAGCAUCAUCGCCUACACCAUGAGCCUGGGC GCCGAGAACAGCGUGGCCUACAGCAACAACAGCAUCGCC AUCCCCACCAACUUCACCAUCAGCGUGACCACCGAGAUUC UGCCCGUGAGCAUGACCAAGACCAGCGUGGACUGCACCA UGUACAUCUGCGGCGACAGCACCGAGUGCAGCAACCUGC UGCUGCAGUACGGCAGCUUCUGCACCCAGCUGAACCGGG CCCUGACCGGCAUCGCCGUGGAGCAGGACAAGAACACCCA GGAGGUGUUCGCCCAGGUGAAGCAGAUCUACAAGACCCC UCCCAUCAAGGACUUCGGCGGCUUCAACUUCAGCCAGAU CCUGCCCGACCCCAGCAAGCCCAGCAAGCGGAGCUUCAUC GAGGACCUGCUGUUCAACAAGGUGACCCUAGCCGACGCC GGCUUCAUCAAGCAGUACGGCGACUGCCUCGGCGACAUA GCCGCCCGGGACCUGAUCUGCGCCCAGAAGUUCAACGGCC UGACCGUGCUGCCUCCCCUGCUGACCGACGAGAUGAUCGC CCAGUACACCAGCGCCCUGUUAGCCGGAACCAUCACCAGC GGCUGGACUUUCGGCGCUGGAGCCGCUCUGCAGAUCCCC UUCGCCAUGCAGAUGGCCUACCGGUUCAACGGCAUCGGC GUGACCCAGAACGUGCUGUACGAGAACCAGAAGCUGAUC GCCAACCAGUUCAACAGCGCCAUCGGCAAGAUCCAGGAC AGCCUGAGCAGCACCGCUAGCGCCCUGGGCAAGCUGCAG GACGUGGUGAACCAGAACGCCCAGGCCCUGAACACCCUG GUGAAGCAGCUGAGCAGCAACUUCGGCGCCAUCAGCAGC GUGCUGAACGACAUCCUGAGCCGGCUGGACAAGGUGGAG GCCGAGGUGCAGAUCGACCGGCUGAUCACUGGCCGGCUG CAGAGCCUGCAGACCUACGUGACCCAGCAGCUGAUCCGG GCCGCCGAGAUUCGGGCCAGCGCCAACCUGGCCGCCACCA AGAUGAGCGAGUGCGUGCUGGGCCAGAGCAAGCGGGUGG ACUUCUGCGGCAAGGGCUACCACCUGAUGAGCUUUCCCC AGAGCGCACCCCACGGAGUGGUGUUCCUGCACGUGACCU ACGUGCCCGCCCAGGAGAAGAACUUCACCACCGCCCCAGC CAUCUGCCACGACGGCAAGGCCCACUUUCCCCGGGAGGGC GUGUUCGUGAGCAACGGCACCCACUGGUUCGUGACCCAG CGGAACUUCUACGAGCCCCAGAUCAUCACCACCGACAACA CCUUCGUGAGCGGCAACUGCGACGUGGUGAUCGGCAUCG UGAACAACACCGUGUACGAUCCCCUGCAGCCCGAGCUGG ACAGCUUCAAGGAGGAGCUGGACAAGUACUUCAAGAAUC ACACCAGCCCCGACGUGGACCUGGGCGACAUCAGCGGCAU CAACGCCAGCGUGGUGAACAUCCAGAAGGAGAUCGAUCG GCUGAACGAGGUGGCCAAGAACCUGAACGAGAGCCUGAU CGACCUGCAGGAGCUGGGCAAGUACGAGCAGUACAUCAA GUGGCCCUGGUACAUCUGGCUGGGCUUCAUCGCCGGCCU GAUCGCCAUCGUGAUGGUGACCAUCAUGCUG	

3′ UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCC CCUUGGGCCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACC CGUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGG C	4
Corresponding amino acid sequence	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVF RSSVLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPF NDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKV CEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVS QPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRD LPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWT AGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKC TLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATR FASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLN DLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTG CVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQ AGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFE LLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKK FLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTN TSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQT RAGCLIGAEHVNNSYECDIPIGAGICASYQTQTVSLRSVASQSII AYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVD CTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQE VFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNK VTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPLLTDE MIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIG VTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVN QNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLI TGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKR VDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAI CHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFVSG NCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLG DISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIK WPWYIWLGFIAGLIAIVMVTIML	63
PolyA tail	100 nt	
SARS-CoV-2 Variant 26		
SEQ ID NO: 87 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.		87
Chemistry	1-methylpseudouridine	
Cap	7mG(5′)ppp(5′)NlmpNp	
5′ UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUA AGACCCCGGCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGUGAGCAGC CAGUGCGUGAACCUGACCACCCGGACCCAGCUGCCACCAG CCUACACCAACAGCUUCACCCGGGGCGUCUACUACCCCGA CAAGGUGUUCCGGAGCAGCGUCCUGCACAGCACCCAGGA CCUGUUCCUGCCCUUCUUCAGCAACGUGACCUGGUUCCAC GCCAUCCACGUGAGCGGCACCAACGGCACCAAGCGGUUCG ACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUCGC CAGCACCGAGAAGAGCAACAUCAUCCGGGGCUGGAUCUU CGGCACCACCCUGGACAGCAAGACCCAGAGCCUGCUGAUC GUGAAUAACGCCACCAACGUGGUGAUCAAGGUGUGCGAG UUCCAGUUCUGCAACGACCCCUUCCUGGGCGUGUACUACC ACAAGAACAACAAGAGCUGGAUGGAGAGCGAGUUCCGGG UGUACAGCAGCGCCAACAACUGCACCUUCGAGUACGUGA GCCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAGGGCA ACUUCAAGAACCUGCGGGAGUUCGUGUUCAAGAACAUCG ACGGCUACUUCAAGAUCUACAGCAAGCACACCCCAAUCA ACCUGGUGCGGGAUCUGCCCCAGGGCUUCUCAGCCCUGG	64

	AGCCCCUGGUGGACCUGCCCAUCGGCAUCAACAUCACCCG GUUCCAGACCCUGCUGGCCCUGCACCGGAGCUACCUGACC CCAGGCGACAGCAGCAGCGGGUGGACAGCAGGCGCGGCU GCUUACUACGUGGGCUACCUGCAGCCCCGGACCUUCCUGC UGAAGUACAACGAGAACGGCACCAUCACCGACGCCGUGG ACUGCGCCCUGGACCCUCUGAGCGAGACCAAGUGCACCCU GAAGAGCUUCACCGUGGAGAAGGGCAUCUACCAGACCAG CAACUUCCGGGUGCAGCCCACCGAGAGCAUCGUGCGGUU CCCCAACAUCACCAACCUGUGCCCCUUCGGCGAGGUGUUC AACGCCACCCGGUUCGCCAGCGUGUACGCCUGGAACCGGA AGCGGAUCAGCAACUGCGUGGCCGACUACAGCGUGCUGU ACAACAGCGCCAGCUUCAGCACCUUCAAGUGCUACGGCG UGAGCCCCACCAAGCUGAACGACCUGUGCUUCACCAACGU GUACGCCGACAGCUUCGUGAUCCGUGGCGACGAGGUGCG GCAGAUCGCACCCGGCCAGACAGGCAAGAUCGCCGACUAC AACUACAAGCUGCCCGACGACUUCACCGGCUGCGUGAUC GCCUGGAACAGCAACAACCUCGACAGCAAGGUGGGCGGC AACUACAACUACCUGUACCGGCUGUUCCGGAAGAGCAAC CUGAAGCCCUUCGAGCGGGACAUCAGCACCGAGAUCUAC CAAGCCGGCUCCACCCCUUGCAACGGCGUGGAGGGCUUCA ACUGCUACUUCCCUCUGCAGAGCUACGGCUUCCAGCCCAC CAACGGCGUGGGCUACCAGCCCUACCGGGUGGUGGUGCU GAGCUUCGAGCUGCUGCACGCCCCAGCCACCGUGUGUGGC CCCAAGAAGAGCACCAACCUGGUGAAGAACAAGUGCGUG AACUUCAACUUCAACGGCCUUACCGGCACCGGCGUGCUG ACCGAGAGCAACAAGAAAUUCCUGCCCUUUCAGCAGUUC GGCCGGGACAUCGCCGACACCACCGACGCUGUGCGGGAUC CCCAGACCCUGGAGAUCCUGGACAUCACCCCUUGCAGCUU CGGCGGCGUGAGCGUGAUCACCCCAGGCACCAACACCAGC AACCAGGUGGCCGUGCUGUACCAGGACGUGAACUGCACC GAGGUGCCCGUGGCCAUCCACGCCGACCAGCUGACACCCA CCUGGCGGGUCUACAGCACCGGCAGCAACGUGUUCCAGA CCCGGGCCGGUUGCCUGAUCGGCGCCGAGCACGUGAACA ACAGCUACGAGUGCGACAUCCCCAUCGGCGCCGGCAUCUG UGCCAGCUACCAGACCCAGACCGUGUCACUGAGGAGCGU GGCCAGCCAGAGCAUCAUCGCCUACACCAUGAGCCUGGGC GCCGAGAACAGCGUGGCCUACAGCAACAACAGCAUCGCC AUCCCCACCAACUUCACCAUCAGCGUGACCACCGAGAUUC UGCCCGUGAGCAUGACCAAGACCAGCGUGGACUGCACCA UGUACAUCUGCGGCGACAGCACCGAGUGCAGCAACCUGC UGCUGCAGUACGGCAGCUUCUGCACCCAGCUGAACCGGG CCCUGACCGGCAUCGCCGUGGAGCAGGACAAGAACACCCA GGAGGUGUUCGCCCAGGUGAAGCAGAUCUACAAGACCCC UCCCAUCAAGGACUUCGGCGGCUUCAACUUCAGCCAGAU CCUGCCCGACCCCAGCAAGCCCAGCAAGCGGAGCUUCAUC GAGGACCUGCUGUUCAACAAGGUGACCCUAGCCGACGCC GGCUUCAUCAAGCAGUACGGCGACUGCCUCGGCGACAUA GCCGCCCGGGACCUGAUCUGCGCCCAGAAGUUCAACGGCC UGACCGUGCUGCCUCCCCUGCUGACCGACGAGAUGAUCGC CCAGUACACCAGCGCCCUGUUAGCCGGAACCAUCACCAGC GGCUGGACUUUCGGCGCUGGAGCCGCUCUGCAGAUCCCC UUCGCCAUGCAGAUGGCCUACCGGUUCAACGGCAUCGGC GUGACCCAGAACGUGCUGUACGAGAACCAGAAGCUGAUC GCCAACCAGUUCAACAGCGCCAUCGGCAAGAUCCAGGAC AGCCUGAGCAGCACCGCUAGCGCCCUGGGCAAGCUGCAG GACGUGGUGAACCAGAACGCCCAGGCCCUGAACACCCUG GUGAAGCAGCUGAGCAGCAACUUCGGCGCCAUCAGCAGC GUGCUGAACGACAUCCUGAGCCGGCUGGACAAGGUGGAG	

	GCCGAGGUGCAGAUCGACCGGCUGAUCACUGGCCGGCUG CAGAGCCUGCAGACCUACGUGACCCAGCAGCUGAUCCGG GCCGCCGAGAUUCGGGCCAGCGCCAACCUGGCCGCCACCA AGAUGAGCGAGUGCGUGCUGGGCCAGAGCAAGCGGGUGG ACUUCUGCGGCAAGGGCUACCACCUGAUGAGCUUUCCCC AGAGCGCACCCCACGGAGUGGUGUUCCUGCACGUGACCU ACGUGCCCGCCCAGGAGAAGAACUUCACCACCGCCCCAGC CAUCUGCCACGACGGCAAGGCCCACUUUCCCCGGGAGGGC GUGUUCGUGAGCAACGGCACCCACUGGUUCGUGACCCAG CGGAACUUCUACGAGCCCCAGAUCAUCACCACCGACAACA CCUUCGUGAGCGGCAACUGCGACGUGGUGAUCGGCAUCG UGAACAACACCGUGUACGAUCCCCUGCAGCCCGAGCUGG ACAGCUUCAAGGAGGAGCUGGACAAGUACUUCAAGAAUC ACACCAGCCCCGACGUGGACCUGGGCGACAUCAGCGGCAU CAACGCCAGCGUGGUGAACAUCCAGAAGGAGAUCGAUCG GCUGAACGAGGUGGCCAAGAACCUGAACGAGAGCCUGAU CGACCUGCAGGAGCUGGGCAAGUACGAGCAGUACAUCAA GUGGCCCUGGUACAUCUGGCUGGGCUUCAUCGCCGGCCU GAUCGCCAUCGUGAUGGUGACCAUCAUGCUGAAGAAGAA GAAGCGGCCACGGAACUCCUACAAGUGCGGCACCAACACC AUGGAGCGGGAGGAGAGCGAGCAGACCAAGAAGCGGGAG AAGAUCCACAUUCCUGAACGGUCCGACGAAGCCCAGCGG GUGUUCAAGAGCAGCAAGACCAGCAGCUGCGACAAGAGC GACACCUGCUUC	
3′ UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCC CCUUGGGCCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACC CGUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGG C	4
Corresponding amino acid sequence	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVF RSSVLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPF NDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKV CEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVS QPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRD LPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWT AGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKC TLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATR FASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLN DLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTG CVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQ AGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFE LLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKK FLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTN TSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQT RAGCLIGAEHVNNSYECDIPIGAGICASYQTQTVSLRSVASQSII AYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVD CTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQE VFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNK VTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPLLTDE MIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIG VTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVN QNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLI TGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKR VDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAI CHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFVSG NCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLG DISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIK WPWYIWLGFIAGLIAIVMVTIMLKKKKRPRNSYKCGTNTMER	65

	EESEQTKKREKIHIPERSDEAQRVFKSSKTSSCDKSDTCF	
PolyA tail	100 nt	
SARS-CoV-2 Variant 27		
SEQ ID NO: 88 consists of from 5′ end to 3′ end: 5′ UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 66, and 3′ UTR SEQ ID NO: 4.		88
Chemistry	1-methylpseudouridine	
Cap	7mG(5′)ppp(5′)NlmpNp	
5′ UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUA AGACCCCGGCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGUGAGCAGC CAGUGCGUGAACCUGACCACCCGGACCCAGCUGCCACCAG CCUACACCAACAGCUUCACCCGGGGCGUCUACUACCCCGA CAAGGUGUUCCGGAGCAGCGUCCUGCACAGCACCCAGGA CCUGUUCCUGCCCUUCUUCAGCAACGUGACCUGGUUCCAC GCCAUCCACGUGAGCGGCACCAACGGCACCAAGCGGUUCG ACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUCGC CAGCACCGAGAAGAGCAACAUCAUCCGGGGCUGGAUCUU CGGCACCACCCUGGACAGCAAGACCCAGAGCCUGCUGAUC GUGAAUAACGCCACCAACGUGGUGAUCAAGGUGUGCGAG UUCCAGUUCUGCAACGACCCCUUCCUGGGCGUGUACUACC ACAAGAACAACAAGAGCUGGAUGGAGAGCGAGUUCCGGG UGUACAGCAGCGCCAACAACUGCACCUUCGAGUACGUGA GCCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAGGGCA ACUUCAAGAACCUGCGGGAGUUCGUGUUCAAGAACAUCG ACGGCUACUUCAAGAUCUACAGCAAGCACACCCCAAUCA ACCUGGUGCGGGAUCUGCCCCAGGGCUUCUCAGCCCUGG AGCCCCUGGUGGACCUGCCCAUCGGCAUCAACAUCACCCG GUUCCAGACCCUGCUGGCCCUGCACCGGAGCUACCUGACC CCAGGCGACAGCAGCAGCGGGUGGACAGCAGGCGCGGCU GCUUACUACGUGGGCUACCUGCAGCCCCGGACCUUCCUGC UGAAGUACAACGAGAACGGCACCAUCACCGACGCCGUGG ACUGCGCCCUGGACCCUCUGAGCGAGACCAAGUGCACCCU GAAGAGCUUCACCGUGGAGAAGGGCAUCUACCAGACCAG CAACUUCCGGGUGCAGCCCACCGAGAGCAUCGUGCGGUU CCCCAACAUCACCAACCUGUGCCCCUUCGGCGAGGUGUUC AACGCCACCCGGUUCGCCAGCGUGUACGCCUGGAACCGGA AGCGGAUCAGCAACUGCGUGGCCGACUACAGCGUGCUGU ACAACAGCGCCAGCUUCAGCACCUUCAAGUGCUACGGCG UGAGCCCCACCAAGCUGAACGACCUGUGCUUCACCAACGU GUACGCCGACAGCUUCGUGAUCCGUGGCGACGAGGUGCG GCAGAUCGCACCCGGCCAGACAGGCAAGAUCGCCGACUAC AACUACAAGCUGCCCGACGACUUCACCGGCUGCGUGAUC GCCUGGAACAGCAACAACCUCGACAGCAAGGUGGGCGGC AACUACAACUACCUGUACCGGCUGUUCCGGAAGAGCAAC CUGAAGCCCUUCGAGCGGGACAUCAGCACCGAGAUCUAC CAAGCCGGCUCCACCCCUUGCAACGGCGUGGAGGGCUUCA ACUGCUACUUCCCUCUGCAGAGCUACGGCUUCCAGCCCAC CAACGGCGUGGGCUACCAGCCCUACCGGGUGGUGGUGCU GAGCUUCGAGCUGCUGCACGCCCCAGCCACCGUGUGUGGC CCCAAGAAGAGCACCAACCUGGUGAAGAACAAGUGCGUG AACUUCAACUUCAACGGCCUUACCGGCACCGGCGUGCUG ACCGAGAGCAACAAGAAAUUCCUGCCCUUUCAGCAGUUC GGCCGGGACAUCGCCGACACCACCGACGCUGUGCGGGAUC CCCAGACCCUGGAGAUCCUGGACAUCACCCCUUGCAGCUU CGGCGGCGUGAGCGUGAUCACCCCAGGCACCAACACCAGC	66

	AACCAGGUGGCCGUGCUGUACCAGGACGUGAACUGCACC GAGGUGCCCGUGGCCAUCCACGCCGACCAGCUGACACCCA CCUGGCGGGUCUACAGCACCGGCAGCAACGUGUUCCAGA CCCGGGCCGGUUGCCUGAUCGGCGCCGAGCACGUGAACA ACAGCUACGAGUGCGACAUCCCCAUCGGCGCCGGCAUCUG UGCCAGCUACCAGACCCAGACCGUGUCACUGAGGAGCGU GGCCAGCCAGAGCAUCAUCGCCUACACCAUGAGCCUGGGC GCCGAGAACAGCGUGGCCUACAGCAACAACAGCAUCGCC AUCCCCACCAACUUCACCAUCAGCGUGACCACCGAGAUUC UGCCCGUGAGCAUGACCAAGACCAGCGUGGACUGCACCA UGUACAUCUGCGGCGACAGCACCGAGUGCAGCAACCUGC UGCUGCAGUACGGCAGCUUCUGCACCCAGCUGAACCGGG CCCUGACCGGCAUCGCCGUGGAGCAGGACAAGAACACCCA GGAGGUGUUCGCCCAGGUGAAGCAGAUCUACAAGACCCC UCCCAUCAAGGACUUCGGCGGCUUCAACUUCAGCCAGAU CCUGCCCGACCCCAGCAAGCCCAGCAAGCGGAGCUUCAUC GAGGACCUGCUGUUCAACAAGGUGACCCUAGCCGACGCC GGCUUCAUCAAGCAGUACGGCGACUGCCUCGGCGACAUA GCCGCCCGGGACCUGAUCUGCGCCCAGAAGUUCAACGGCC UGACCGUGCUGCCUCCCCUGCUGACCGACGAGAUGAUCGC CCAGUACACCAGCGCCCUGUUAGCCGGAACCAUCACCAGC GGCUGGACUUUCGGCGCUGGAGCCGCUCUGCAGAUCCCC UUCGCCAUGCAGAUGGCCUACCGGUUCAACGGCAUCGGC GUGACCCAGAACGUGCUGUACGAGAACCAGAAGCUGAUC GCCAACCAGUUCAACAGCGCCAUCGGCAAGAUCCAGGAC AGCCUGAGCAGCACCGCUAGCGCCCUGGGCAAGCUGCAG GACGUGGUGAACCAGAACGCCCAGGCCCUGAACACCCUG GUGAAGCAGCUGAGCAGCAACUUCGGCGCCAUCAGCAGC GUGCUGAACGACAUCCUGAGCCGGCUGGACAAGGUGGAG GCCGAGGUGCAGAUCGACCGGCUGAUCACUGGCCGGCUG CAGAGCCUGCAGACCUACGUGACCCAGCAGCUGAUCCGG GCCGCCGAGAUUCGGGCCAGCGCCAACCUGGCCGCCACCA AGAUGAGCGAGUGCGUGCUGGGCCAGAGCAAGCGGGUGG ACUUCUGCGGCAAGGGCUACCACCUGAUGAGCUUUCCCC AGAGCGCACCCCACGGAGUGGUGUUCCUGCACGUGACCU ACGUGCCCGCCCAGGAGAAGAACUUCACCACCGCCCCAGC CAUCUGCCACGACGGCAAGGCCCACUUUCCCCGGGAGGGC GUGUUCGUGAGCAACGGCACCCACUGGUUCGUGACCCAG CGGAACUUCUACGAGCCCCAGAUCAUCACCACCGACAACA CCUUCGUGAGCGGCAACUGCGACGUGGUGAUCGGCAUCG UGAACAACACCGUGUACGAUCCCCUGCAGCCCGAGCUGG ACAGCUUCAAGGAGGAGCUGGACAAGUACUUCAAGAAUC ACACCAGCCCCGACGUGGACCUGGGCGACAUCAGCGGCAU CAACGCCAGCGUGGUGAACAUCCAGAAGGAGAUCGAUCG GCUGAACGAGGUGGCCAAGAACCUGAACGAGAGCCUGAU CGACCUGCAGGAGCUGGGCAAGUACGAGCAGUACAUCAA GUGGCCCUGGUACAUCUGGCUGGGCUUCAUCGCCGGCCU GAUCGCCAUCGUGAUGGUGACCAUCAUGCUGAAGCGGCA GUACAAGGACAUGAUGAGCGAGGGAGGACCACCUGGCGC UGAGCCACAG	
3′ UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCC CCUUGGGCCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACC CGUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGG C	4
Corresponding amino acid sequence	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVF RSSVLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPF NDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKV CEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVS	67

	QPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRD LPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWT AGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKC TLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATR FASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLN DLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTG CVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQ AGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFE LLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKK FLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTN TSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQT RAGCLIGAEHVNNSYECDIPIGAGICASYQTQTVSLRSVASQSII AYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVD CTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQE VFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNK VTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPLLTDE MIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIG VTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVN QNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLI TGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKR VDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAI CHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFVSG NCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLG DISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIK WPWYIWLGFIAGLIAIVMVTIMLKRQYKDMMSEGGPPGAEPQ	
PolyA tail	100 nt	
SARS-CoV-2 Variant 28		
SEQ ID NO: 89 consists of from 5′ end to 3′ end: 5′ UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 68, and 3′ UTR SEQ ID NO: 4.		89
Chemistry	1-methylpseudouridine	
Cap	7mG(5′)ppp(5′)NlmpNp	
5′ UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUA AGACCCCGGCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGUGAGCAGC CAGUGCGUGAACCUGACCACCCGGACCCAGCUGCCACCAG CCUACACCAACAGCUUCACCCGGGGCGUCUACUACCCCGA CAAGGUGUUCCGGAGCAGCGUCCUGCACAGCACCCAGGA CCUGUUCCUGCCCUUCUUCAGCAACGUGACCUGGUUCCAC GCCAUCCACGUGAGCGGCACCAACGGCACCAAGCGGUUCG ACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUCGC CAGCACCGAGAAGAGCAACAUCAUCCGGGGCUGGAUCUU CGGCACCACCCUGGACAGCAAGACCCAGAGCCUGCUGAUC GUGAAUAACGCCACCAACGUGGUGAUCAAGGUGUGCGAG UUCCAGUUCUGCAACGACCCCUUCCUGGGCGUGUACUACC ACAAGAACAACAAGAGCUGGAUGGAGAGCGAGUUCCGGG UGUACAGCAGCGCCAACAACUGCACCUUCGAGUACGUGA GCCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAGGGCA ACUUCAAGAACCUGCGGGAGUUCGUGUUCAAGAACAUCG ACGGCUACUUCAAGAUCUACAGCAAGCACACCCCAAUCA ACCUGGUGCGGGAUCUGCCCCAGGGCUUCUCAGCCCUGG AGCCCCUGGUGGACCUGCCCAUCGGCAUCAACAUCACCCG GUUCCAGACCCUGCUGGCCCUGCACCGGAGCUACCUGACC CCAGGCGACAGCAGCAGCGGGUGGACAGCAGGCGCGGCU GCUUACUACGUGGGCUACCUGCAGCCCCGGACCUUCCUGC UGAAGUACAACGAGAACGGCACCAUCACCGACGCCGUGG	68

	ACUGCGCCCUGGACCCUCUGAGCGAGACCAAGUGCACCCU GAAGAGCUUCACCGUGGAGAAGGGCAUCUACCAGACCAG CAACUUCCGGGUGCAGCCCACCGAGAGCAUCGUGCGGUU CCCCAACAUCACCAACCUGUGCCCCUUCGGCGAGGUGUUC AACGCCACCCGGUUCGCCAGCGUGUACGCCUGGAACCGGA AGCGGAUCAGCAACUGCGUGGCCGACUACAGCGUGCUGU ACAACAGCGCCAGCUUCAGCACCUUCAAGUGCUACGGCG UGAGCCCCACCAAGCUGAACGACCUGUGCUUCACCAACGU GUACGCCGACAGCUUCGUGAUCCGUGGCGACGAGGUGCG GCAGAUCGCACCCGGCCAGACAGGCAAGAUCGCCGACUAC AACUACAAGCUGCCCGACGACUUCACCGGCUGCGUGAUC GCCUGGAACAGCAACAACCUCGACAGCAAGGUGGGCGGC AACUACAACUACCUGUACCGGCUGUUCCGGAAGAGCAAC CUGAAGCCCUUCGAGCGGGACAUCAGCACCGAGAUCUAC CAAGCCGGCUCCACCCCUUGCAACGGCGUGGAGGGCUUCA ACUGCUACUUCCCUCUGCAGAGCUACGGCUUCCAGCCCAC CAACGGCGUGGGCUACCAGCCCUACCGGGUGGUGGUGCU GAGCUUCGAGCUGCUGCACGCCCCAGCCACCGUGUGUGGC CCCAAGAAGAGCACCAACCUGGUGAAGAACAAGUGCGUG AACUUCAACUUCAACGGCCUUACCGGCACCGGCGUGCUG ACCGAGAGCAACAAGAAAUUCCUGCCCUUUCAGCAGUUC GGCCGGGACAUCGCCGACACCACCGACGCUGUGCGGGAUC CCCAGACCCUGGAGAUCCUGGACAUCACCCCUUGCAGCUU CGGCGGCGUGAGCGUGAUCACCCCAGGCACCAACACCAGC AACCAGGUGGCCGUGCUGUACCAGGACGUGAACUGCACC GAGGUGCCCGUGGCCAUCCACGCCGACCAGCUGACACCCA CCUGGCGGGUCUACAGCACCGGCAGCAACGUGUUCCAGA CCCGGGCCGGUUGCCUGAUCGGCGCCGAGCACGUGAACA ACAGCUACGAGUGCGACAUCCCCAUCGGCGCCGGCAUCUG UGCCAGCUACCAGACCCAGACCGUGUCACUGAGGAGCGU GGCCAGCCAGAGCAUCAUCGCCUACACCAUGAGCCUGGGC GCCGAGAACAGCGUGGCCUACAGCAACAACAGCAUCGCC AUCCCCACCAACUUCACCAUCAGCGUGACCACCGAGAUUC UGCCCGUGAGCAUGACCAAGACCAGCGUGGACUGCACCA UGUACAUCUGCGGCGACAGCACCGAGUGCAGCAACCUGC UGCUGCAGUACGGCAGCUUCUGCACCCAGCUGAACCGGG CCCUGACCGGCAUCGCCGUGGAGCAGGACAAGAACACCCA GGAGGUGUUCGCCCAGGUGAAGCAGAUCUACAAGACCCC UCCCAUCAAGGACUUCGGCGGCUUCAACUUCAGCCAGAU CCUGCCCGACCCCAGCAAGCCCAGCAAGCGGAGCUUCAUC GAGGACCUGCUGUUCAACAAGGUGACCCUAGCCGACGCC GGCUUCAUCAAGCAGUACGGCGACUGCCUCGGCGACAUA GCCGCCCGGGACCUGAUCUGCGCCCAGAAGUUCAACGGCC UGACCGUGCUGCCUCCCCUGCUGACCGACGAGAUGAUCGC CCAGUACACCAGCGCCCUGUUAGCCGGAACCAUCACCAGC GGCUGGACUUUCGGCGCUGGAGCCGCUCUGCAGAUCCCC UUCGCCAUGCAGAUGGCCUACCGGUUCAACGGCAUCGGC GUGACCCAGAACGUGCUGUACGAGAACCAGAAGCUGAUC GCCAACCAGUUCAACAGCGCCAUCGGCAAGAUCCAGGAC AGCCUGAGCAGCACCGCUAGCGCCCUGGGCAAGCUGCAG GACGUGGUGAACCAGAACGCCCAGGCCCUGAACACCCUG GUGAAGCAGCUGAGCAGCAACUUCGGCGCCAUCAGCAGC GUGCUGAACGACAUCCUGAGCCGGCUGGACCCCCCCGAGG CCGAGGUGCAGAUCGACCGGCUGAUCACUGGCCGGCUGC AGAGCCUGCAGACCUACGUGACCCAGCAGCUGAUCCGGG CCGCCGAGAUUCGGGCCAGCGCCAACCUGGCCGCCACCAA GAUGAGCGAGUGCGUGCUGGGCCAGAGCAAGCGGGUGGA CUUCUGCGGCAAGGGCUACCACCUGAUGAGCUUUCCCCA	

	GAGCGCACCCCACGGAGUGGUGUUCCUGCACGUGACCUA CGUGCCCGCCCAGGAGAAGAACUUCACCACCGCCCCAGCC AUCUGCCACGACGGCAAGGCCCACUUUCCCCGGGAGGGCG UGUUCGUGAGCAACGGCACCCACUGGUUCGUGACCCAGC GGAACUUCUACGAGCCCCAGAUCAUCACCACCGACAACAC CUUCGUGAGCGGCAACUGCGACGUGGUGAUCGGCAUCGU GAACAACACCGUGUACGAUCCCCUGCAGCCCGAGCUGGAC AGCUUCAAGGAGGAGCUGGACAAGUACUUCAAGAAUCAC ACCAGCCCCGACGUGGACCUGGGCGACAUCAGCGGCAUCA ACGCCAGCGUGGUGAACAUCCAGAAGGAGAUCGAUCGGC UGAACGAGGUGGCCAAGAACCUGAACGAGAGCCUGAUCG ACCUGCAGGAGCUGGGCAAGUACGAGCAGUACAUCAAGU GGCCCUGGUACAUCUGGCUGGGCUUCAUCGCCGGCCUGA UCGCCAUCGUGAUGGUGACCAUCAUGCUG	
3′ UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCC CCUUGGGCCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACC CGUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGG C	4
Corresponding amino acid sequence	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVF RSSVLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPF NDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKV CEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVS QPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRD LPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWT AGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKC TLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATR FASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLN DLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTG CVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQ AGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFE LLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKK FLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTN TSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQT RAGCLIGAEHVNNSYECDIPIGAGICASYQTQTVSLRSVASQSII AYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVD CTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQE VFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNK VTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPLLTDE MIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIG VTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVN QNAQALNTLVKQLSSNFGAISSVLNDILSRLDPPEAEVQIDRLI TGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKR VDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAI CHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFVSG NCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLG DISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIK WPWYIWLGFIAGLIAIVMVTIML	69
PolyA tail	100 nt	
SARS-CoV-2 Variant 29		
SEQ ID NO: 90 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.		90
Chemistry	1-methylpseudouridine	
Cap	7mG(5′)ppp(5′)NlmpNp	
5′ UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUA AGACCCCGGCGCCGCCACC	2

ORF of mRNA Construct (excluding the stop codon)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGUGAGCAGC CAGUGCGUGAACCUGACCACCCGGACCCAGCUGCCACCAG CCUACACCAACAGCUUCACCCGGGGCGUCUACUACCCCGA CAAGGUGUUCCGGAGCAGCGUCCUGCACAGCACCCAGGA CCUGUUCCUGCCCUUCUUCAGCAACGUGACCUGGUUCCAC GCCAUCCACGUGAGCGGCACCAACGGCACCAAGCGGUUCG ACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUCGC CAGCACCGAGAAGAGCAACAUCAUCCGGGGCUGGAUCUU CGGCACCACCCUGGACAGCAAGACCCAGAGCCUGCUGAUC GUGAAUAACGCCACCAACGUGGUGAUCAAGGUGUGCGAG UUCCAGUUCUGCAACGACCCCUUCCUGGGCGUGUACUACC ACAAGAACAACAAGAGCUGGAUGGAGAGCGAGUUCCGGG UGUACAGCAGCGCCAACAACUGCACCUUCGAGUACGUGA GCCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAGGGCA ACUUCAAGAACCUGCGGGAGUUCGUGUUCAAGAACAUCG ACGGCUACUUCAAGAUCUACAGCAAGCACACCCCAAUCA ACCUGGUGCGGGAUCUGCCCCAGGGCUUCUCAGCCCUGG AGCCCCUGGUGGACCUGCCCAUCGGCAUCAACAUCACCCG GUUCCAGACCCUGCUGGCCCUGCACCGGAGCUACCUGACC CCAGGCGACAGCAGCAGCGGGUGGACAGCAGGCGCGGCU GCUUACUACGUGGGCUACCUGCAGCCCCGGACCUUCCUGC UGAAGUACAACGAGAACGGCACCAUCACCGACGCCGUGG ACUGCGCCCUGGACCCUCUGAGCGAGACCAAGUGCACCCU GAAGAGCUUCACCGUGGAGAAGGGCAUCUACCAGACCAG CAACUUCCGGGUGCAGCCCACCGAGAGCAUCGUGCGGUU CCCCAACAUCACCAACCUGUGCCCCUUCGGCGAGGUGUUC AACGCCACCCGGUUCGCCAGCGUGUACGCCUGGAACCGGA AGCGGAUCAGCAACUGCGUGGCCGACUACAGCGUGCUGU ACAACAGCGCCAGCUUCAGCACCUUCAAGUGCUACGGCG UGAGCCCCACCAAGCUGAACGACCUGUGCUUCACCAACGU GUACGCCGACAGCUUCGUGAUCCGUGGCGACGAGGUGCG GCAGAUCGCACCCGGCCAGACAGGCAAGAUCGCCGACUAC AACUACAAGCUGCCCGACGACUUCACCGGCUGCGUGAUC GCCUGGAACAGCAACAACCUCGACAGCAAGGUGGGCGGC AACUACAACUACCUGUACCGGCUGUUCCGGAAGAGCAAC CUGAAGCCCUUCGAGCGGGACAUCAGCACCGAGAUCUAC CAAGCCGGCUCCACCCCUUGCAACGGCGUGGAGGGCUUCA ACUGCUACUUCCCUCUGCAGAGCUACGGCUUCCAGCCCAC CAACGGCGUGGGCUACCAGCCCUACCGGGUGGUGGUGCU GAGCUUCGAGCUGCUGCACGCCCCAGCCACCGUGUGUGGC CCCAAGAAGAGCACCAACCUGGUGAAGAACAAGUGCGUG AACUUCAACUUCAACGGCCUUACCGGCACCGGCGUGCUG ACCGAGAGCAACAAGAAAUUCCUGCCCUUUCAGCAGUUC GGCCGGGACAUCGCCGACACCACCGACGCUGUGCGGGAUC CCCAGACCCUGGAGAUCCUGGACAUCACCCCUUGCAGCUU CGGCGGCGUGAGCGUGAUCACCCCAGGCACCAACACCAGC AACCAGGUGGCCGUGCUGUACCAGGACGUGAACUGCACC GAGGUGCCCGUGGCCAUCCACGCCGACCAGCUGACACCCA CCUGGCGGGUCUACAGCACCGGCAGCAACGUGUUCCAGA CCCGGGCCGGUUGCCUGAUCGGCGCCGAGCACGUGAACA ACAGCUACGAGUGCGACAUCCCCAUCGGCGCCGGCAUCUG UGCCAGCUACCAGACCCAGACCGUGUCACUGAGGAGCGU GGCCAGCCAGAGCAUCAUCGCCUACACCAUGAGCCUGGGC GCCGAGAACAGCGUGGCCUACAGCAACAACAGCAUCGCC AUCCCCACCAACUUCACCAUCAGCGUGACCACCGAGAUUC UGCCCGUGAGCAUGACCAAGACCAGCGUGGACUGCACCA UGUACAUCUGCGGCGACAGCACCGAGUGCAGCAACCUGC UGCUGCAGUACGGCAGCUUCUGCACCCAGCUGAACCGGG	70

	CCCUGACCGGCAUCGCCGUGGAGCAGGACAAGAACACCCA GGAGGUGUUCGCCCAGGUGAAGCAGAUCUACAAGACCCC UCCCAUCAAGGACUUCGGCGGCUUCAACUUCAGCCAGAU CCUGCCCGACCCCAGCAAGCCCAGCAAGCGGAGCUUCAUC GAGGACCUGCUGUUCAACAAGGUGACCCUAGCCGACGCC GGCUUCAUCAAGCAGUACGGCGACUGCCUCGGCGACAUA GCCGCCCGGGACCUGAUCUGCGCCCAGAAGUUCAACGGCC UGACCGUGCUGCCUCCCCUGCUGACCGACGAGAUGAUCGC CCAGUACACCAGCGCCCUGUUAGCCGGAACCAUCACCAGC GGCUGGACUUUCGGCGCUGGAGCCGCUCUGCAGAUCCCC UUCGCCAUGCAGAUGGCCUACCGGUUCAACGGCAUCGGC GUGACCCAGAACGUGCUGUACGAGAACCAGAAGCUGAUC GCCAACCAGUUCAACAGCGCCAUCGGCAAGAUCCAGGAC AGCCUGAGCAGCACCGCUAGCGCCCUGGGCAAGCUGCAG GACGUGGUGAACCAGAACGCCCAGGCCCUGAACACCCUG GUGAAGCAGCUGAGCAGCAACUUCGGCGCCAUCAGCAGC GUGCUGAACGACAUCCUGAGCCGGCUGGACCCCCCCGAGG CCGAGGUGCAGAUCGACCGGCUGAUCACUGGCCGGCUGC AGAGCCUGCAGACCUACGUGACCCAGCAGCUGAUCCGGG CCGCCGAGAUUCGGGCCAGCGCCAACCUGGCCGCCACCAA GAUGAGCGAGUGCGUGCUGGGCCAGAGCAAGCGGGUGGA CUUCUGCGGCAAGGGCUACCACCUGAUGAGCUUUCCCCA GAGCGCACCCCACGGAGUGGUGUUCCUGCACGUGACCUA CGUGCCCGCCCAGGAGAAGAACUUCACCACCGCCCCAGCC AUCUGCCACGACGGCAAGGCCCACUUUCCCCGGGAGGGCG UGUUCGUGAGCAACGGCACCCACUGGUUCGUGACCCAGC GGAACUUCUACGAGCCCCAGAUCAUCACCACCGACAACAC CUUCGUGAGCGGCAACUGCGACGUGGUGAUCGGCAUCGU GAACAACACCGUGUACGAUCCCCUGCAGCCCGAGCUGGAC AGCUUCAAGGAGGAGCUGGACAAGUACUUCAAGAAUCAC ACCAGCCCCGACGUGGACCUGGGCGACAUCAGCGGCAUCA ACGCCAGCGUGGUGAACAUCCAGAAGGAGAUCGAUCGGC UGAACGAGGUGGCCAAGAACCUGAACGAGAGCCUGAUCG ACCUGCAGGAGCUGGGCAAGUACGAGCAGUACAUCAAGU GGCCCUGGUACAUCUGGCUGGGCUUCAUCGCCGGCCUGA UCGCCAUCGUGAUGGUGACCAUCAUGCUGAAGAAGAAGA AGCGGCCACGGAACUCCUACAAGUGCGGCACCAACACCAU GGAGCGGGAGGAGAGCGAGCAGACCAAGAAGCGGGAGAA GAUCCACAUUCCUGAACGGUCCGACGAAGCCCAGCGGGU GUUCAAGAGCAGCAAGACCAGCAGCUGCGACAAGAGCGA CACCUGCUUC	
3′ UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCC CCUUGGGCCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACC CGUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGG C	4
Corresponding amino acid sequence	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVF RSSVLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPF NDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKV CEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVS QPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRD LPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWT AGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKC TLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATR FASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLN DLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTG CVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQ AGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFE LLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKK	71

	FLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTN TSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQT RAGCLIGAEHVNNSYECDIPIGAGICASYQTQTVSLRSVASQSII AYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVD CTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQE VFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNK VTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPLLTDE MIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIG VTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVN QNAQALNTLVKQLSSNFGAISSVLNDILSRLDPPEAEVQIDRLI TGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKR VDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAI CHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFVSG NCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLG DISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIK WPWYIWLGFIAGLIAIVMVTIMLKKKKRPRNSYKCGTNTMER EESEQTKKREKIHIPERSDEAQRVFKSSKTSSCDKSDTCF	
PolyA tail	100 nt	
SARS-CoV-2 Variant 30		
SEQ ID NO: 91 consists of from 5′ end to 3′ end: 5′ UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 72, and 3′ UTR SEQ ID NO: 4.		91
Chemistry	1-methylpseudouridine	
Cap	7mG(5′)ppp(5′)NlmpNp	
5′ UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUA AGACCCCGGCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGUGAGCAGC CAGUGCGUGAACCUGACCACCCGGACCCAGCUGCCACCAG CCUACACCAACAGCUUCACCCGGGGCGUCUACUACCCCGA CAAGGUGUUCCGGAGCAGCGUCCUGCACAGCACCCAGGA CCUGUUCCUGCCCUUCUUCAGCAACGUGACCUGGUUCCAC GCCAUCCACGUGAGCGGCACCAACGGCACCAAGCGGUUCG ACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUCGC CAGCACCGAGAAGAGCAACAUCAUCCGGGGCUGGAUCUU CGGCACCACCCUGGACAGCAAGACCCAGAGCCUGCUGAUC GUGAAUAACGCCACCAACGUGGUGAUCAAGGUGUGCGAG UUCCAGUUCUGCAACGACCCCUUCCUGGGCGUGUACUACC ACAAGAACAACAAGAGCUGGAUGGAGAGCGAGUUCCGGG UGUACAGCAGCGCCAACAACUGCACCUUCGAGUACGUGA GCCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAGGGCA ACUUCAAGAACCUGCGGGAGUUCGUGUUCAAGAACAUCG ACGGCUACUUCAAGAUCUACAGCAAGCACACCCCAAUCA ACCUGGUGCGGGAUCUGCCCCAGGGCUUCUCAGCCCUGG AGCCCCUGGUGGACCUGCCCAUCGGCAUCAACAUCACCCG GUUCCAGACCCUGCUGGCCCUGCACCGGAGCUACCUGACC CCAGGCGACAGCAGCAGCGGGUGGACAGCAGGCGCGGCU GCUUACUACGUGGGCUACCUGCAGCCCCGGACCUUCCUGC UGAAGUACAACGAGAACGGCACCAUCACCGACGCCGUGG ACUGCGCCCUGGACCCUCUGAGCGAGACCAAGUGCACCCU GAAGAGCUUCACCGUGGAGAAGGGCAUCUACCAGACCAG CAACUUCCGGGUGCAGCCCACCGAGAGCAUCGUGCGGUU CCCCAACAUCACCAACCUGUGCCCCUUCGGCGAGGUGUUC AACGCCACCCGGUUCGCCAGCGUGUACGCCUGGAACCGGA AGCGGAUCAGCAACUGCGUGGCCGACUACAGCGUGCUGU ACAACAGCGCCAGCUUCAGCACCUUCAAGUGCUACGGCG UGAGCCCCACCAAGCUGAACGACCUGUGCUUCACCAACGU	72

	GUACGCCGACAGCUUCGUGAUCCGUGGCGACGAGGUGCG GCAGAUCGCACCCGGCCAGACAGGCAAGAUCGCCGACUAC AACUACAAGCUGCCCGACGACUUCACCGGCUGCGUGAUC GCCUGGAACAGCAACAACCUCGACAGCAAGGUGGGCGGC AACUACAACUACCUGUACCGGCUGUUCCGGAAGAGCAAC CUGAAGCCCUUCGAGCGGGACAUCAGCACCGAGAUCUAC CAAGCCGGCUCCACCCCUUGCAACGGCGUGGAGGGCUUCA ACUGCUACUUCCCUCUGCAGAGCUACGGCUUCCAGCCCAC CAACGGCGUGGGCUACCAGCCCUACCGGGUGGUGGUGCU GAGCUUCGAGCUGCUGCACGCCCCAGCCACCGUGUGUGGC CCCAAGAAGAGCACCAACCUGGUGAAGAACAAGUGCGUG AACUUCAACUUCAACGGCCUUACCGGCACCGGCGUGCUG ACCGAGAGCAACAAGAAAUUCCUGCCCUUUCAGCAGUUC GGCCGGGACAUCGCCGACACCACCGACGCUGUGCGGGAUC CCCAGACCCUGGAGAUCCUGGACAUCACCCCUUGCAGCUU CGGCGGCGUGAGCGUGAUCACCCCAGGCACCAACACCAGC AACCAGGUGGCCGUGCUGUACCAGGACGUGAACUGCACC GAGGUGCCCGUGGCCAUCCACGCCGACCAGCUGACACCCA CCUGGCGGGUCUACAGCACCGGCAGCAACGUGUUCCAGA CCCGGGCCGGUUGCCUGAUCGGCGCCGAGCACGUGAACA ACAGCUACGAGUGCGACAUCCCCAUCGGCGCCGGCAUCUG UGCCAGCUACCAGACCCAGACCGUGUCACUGAGGAGCGU GGCCAGCCAGAGCAUCAUCGCCUACACCAUGAGCCUGGGC GCCGAGAACAGCGUGGCCUACAGCAACAACAGCAUCGCC AUCCCCACCAACUUCACCAUCAGCGUGACCACCGAGAUUC UGCCCGUGAGCAUGACCAAGACCAGCGUGGACUGCACCA UGUACAUCUGCGGCGACAGCACCGAGUGCAGCAACCUGC UGCUGCAGUACGGCAGCUUCUGCACCCAGCUGAACCGGG CCCUGACCGGCAUCGCCGUGGAGCAGGACAAGAACACCCA GGAGGUGUUCGCCCAGGUGAAGCAGAUCUACAAGACCCC UCCCAUCAAGGACUUCGGCGGCUUCAACUUCAGCCAGAU CCUGCCCGACCCCAGCAAGCCCAGCAAGCGGAGCUUCAUC GAGGACCUGCUGUUCAACAAGGUGACCCUAGCCGACGCC GGCUUCAUCAAGCAGUACGGCGACUGCCUCGGCGACAUA GCCGCCCGGGACCUGAUCUGCGCCCAGAAGUUCAACGGCC UGACCGUGCUGCCUCCCCUGCUGACCGACGAGAUGAUCGC CCAGUACACCAGCGCCCUGUUAGCCGGAACCAUCACCAGC GGCUGGACUUUCGGCGCUGGAGCCGCUCUGCAGAUCCCC UUCGCCAUGCAGAUGGCCUACCGGUUCAACGGCAUCGGC GUGACCCAGAACGUGCUGUACGAGAACCAGAAGCUGAUC GCCAACCAGUUCAACAGCGCCAUCGGCAAGAUCCAGGAC AGCCUGAGCAGCACCGCUAGCGCCCUGGGCAAGCUGCAG GACGUGGUGAACCAGAACGCCCAGGCCCUGAACACCCUG GUGAAGCAGCUGAGCAGCAACUUCGGCGCCAUCAGCAGC GUGCUGAACGACAUCCUGAGCCGGCUGGACCCCCCCGAGG CCGAGGUGCAGAUCGACCGGCUGAUCACUGGCCGGCUGC AGAGCCUGCAGACCUACGUGACCCAGCAGCUGAUCCGGG CCGCCGAGAUUCGGGCCAGCGCCAACCUGGCCGCCACCAA GAUGAGCGAGUGCGUGCUGGGCCAGAGCAAGCGGGUGGA CUUCUGCGGCAAGGGCUACCACCUGAUGAGCUUUCCCCA GAGCGCACCCCACGGAGUGGUGUUCCUGCACGUGACCUA CGUGCCCGCCCAGGAGAAGAACUUCACCACCGCCCCAGCC AUCUGCCACGACGGCAAGGCCCACUUUCCCCGGGAGGGCG UGUUCGUGAGCAACGGCACCCACUGGUUCGUGACCCAGC GGAACUUCUACGAGCCCCAGAUCAUCACCACCGACAACAC CUUCGUGAGCGGCAACUGCGACGUGGUGAUCGGCAUCGU GAACAACACCGUGUACGAUCCCCUGCAGCCCGAGCUGGAC AGCUUCAAGGAGGAGCUGGACAAGUACUUCAAGAAUCAC	

	ACCAGCCCCGACGUGGACCUGGGCGACAUCAGCGGCAUCA ACGCCAGCGUGGUGAACAUCCAGAAGGAGAUCGAUCGGC UGAACGAGGUGGCCAAGAACCUGAACGAGAGCCUGAUCG ACCUGCAGGAGCUGGGCAAGUACGAGCAGUACAUCAAGU GGCCCUGGUACAUCUGGCUGGGCUUCAUCGCCGGCCUGA UCGCCAUCGUGAUGGUGACCAUCAUGCUGAAGCGGCAGU ACAAGGACAUGAUGAGCGAGGGAGGACCACCUGGCGCUG AGCCACAG	
3′ UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCC CCUUGGGCCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACC CGUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGG C	4
Corresponding amino acid sequence	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVF RSSVLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPF NDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKV CEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVS QPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRD LPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWT AGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKC TLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATR FASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLN DLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTG CVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQ AGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFE LLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKK FLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTN TSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQT RAGCLIGAEHVNNSYECDIPIGAGICASYQTQTVSLRSVASQSII AYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVD CTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQE VFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNK VTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPLLTDE MIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIG VTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVN QNAQALNTLVKQLSSNFGAISSVLNDILSRLDPPEAEVQIDRLI TGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKR VDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAI CHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFVSG NCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLG DISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIK WPWYIWLGFIAGLIAIVMVTIMLKRQYKDMMSEGGPPGAEPQ	73
PolyA tail	100 nt	
SARS-CoV-2 Variant 31		
SEQ ID NO: 92 consists of from 5′ end to 3′ end: 5′ UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 74, and 3′ UTR SEQ ID NO: 4.		92
Chemistry	1-methylpseudouridine	
Cap	7mG(5′)ppp(5′)NlmpNp	
5′ UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUA AGACCCCGGCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGUGAGCAGC CAGUGCGUGAACCUGACCACCCGGACCCAGCUGCCACCAG CCUACACCAACAGCUUCACCCGGGGCGUCUACUACCCCGA CAAGGUGUUCCGGAGCAGCGUCCUGCACAGCACCCAGGA CCUGUUCCUGCCCUUCUUCAGCAACGUGACCUGGUUCCAC GCCAUCCACGUGAGCGGCACCAACGGCACCAAGCGGUUCG	74

	ACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUCGC CAGCACCGAGAAGAGCAACAUCAUCCGGGGCUGGAUCUU CGGCACCACCCUGGACAGCAAGACCCAGAGCCUGCUGAUC GUGAAUAACGCCACCAACGUGGUGAUCAAGGUGUGCGAG UUCCAGUUCUGCAACGACCCCUUCCUGGGCGUGUACUACC ACAAGAACAACAAGAGCUGGAUGGAGAGCGAGUUCCGGG UGUACAGCAGCGCCAACAACUGCACCUUCGAGUACGUGA GCCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAGGGCA ACUUCAAGAACCUGCGGGAGUUCGUGUUCAAGAACAUCG ACGGCUACUUCAAGAUCUACAGCAAGCACACCCCAAUCA ACCUGGUGCGGGAUCUGCCCCAGGGCUUCUCAGCCCUGG AGCCCCUGGUGGACCUGCCCAUCGGCAUCAACAUCACCCG GUUCCAGACCCUGCUGGCCCUGCACCGGAGCUACCUGACC CCAGGCGACAGCAGCAGCGGGUGGACAGCAGGCGCGGCU GCUUACUACGUGGGCUACCUGCAGCCCCGGACCUUCCUGC UGAAGUACAACGAGAACGGCACCAUCACCGACGCCGUGG ACUGCGCCCUGGACCCUCUGAGCGAGACCAAGUGCACCCU GAAGAGCUUCACCGUGGAGAAGGGCAUCUACCAGACCAG CAACUUCCGGGUGCAGCCCACCGAGAGCAUCGUGCGGUU CCCCAACAUCACCAACCUGUGCCCCUUCGGCGAGGUGUUC AACGCCACCCGGUUCGCCAGCGUGUACGCCUGGAACCGGA AGCGGAUCAGCAACUGCGUGGCCGACUACAGCGUGCUGU ACAACAGCGCCAGCUUCAGCACCUUCAAGUGCUACGGCG UGAGCCCCACCAAGCUGAACGACCUGUGCUUCACCAACGU GUACGCCGACAGCUUCGUGAUCCGUGGCGACGAGGUGCG GCAGAUCGCACCCGGCCAGACAGGCAAGAUCGCCGACUAC AACUACAAGCUGCCCGACGACUUCACCGGCUGCGUGAUC GCCUGGAACAGCAACAACCUCGACAGCAAGGUGGGCGGC AACUACAACUACCUGUACCGGCUGUUCCGGAAGAGCAAC CUGAAGCCCUUCGAGCGGGACAUCAGCACCGAGAUCUAC CAAGCCGGCUCCACCCCUUGCAACGGCGUGGAGGGCUUCA ACUGCUACUUCCCUCUGCAGAGCUACGGCUUCCAGCCCAC CAACGGCGUGGGCUACCAGCCCUACCGGGUGGUGGUGCU GAGCUUCGAGCUGCUGCACGCCCCAGCCACCGUGUGUGGC CCCAAGAAGAGCACCAACCUGGUGAAGAACAAGUGCGUG AACUUCAACUUCAACGGCCUUACCGGCACCGGCGUGCUG ACCGAGAGCAACAAGAAAUUCCUGCCCUUUCAGCAGUUC GGCCGGGACAUCGCCGACACCACCGACGCUGUGCGGGAUC CCCAGACCCUGGAGAUCCUGGACAUCACCCCUUGCAGCUU CGGCGGCGUGAGCGUGAUCACCCCAGGCACCAACACCAGC AACCAGGUGGCCGUGCUGUACCAGGACGUGAACUGCACC GAGGUGCCCGUGGCCAUCCACGCCGACCAGCUGACACCCA CCUGGCGGGUCUACAGCACCGGCAGCAACGUGUUCCAGA CCCGGGCCGGUUGCCUGAUCGGCGCCGAGCACGUGAACA ACAGCUACGAGUGCGACAUCCCCAUCGGCGCCGGCAUCUG UGCCAGCUACCAGACCCAGACCGUGUCACUGAGGAGCGU GGCCAGCCAGAGCAUCAUCGCCUACACCAUGAGCCUGGGC GCCGAGAACAGCGUGGCCUACAGCAACAACAGCAUCGCC AUCCCCACCAACUUCACCAUCAGCGUGACCACCGAGAUUC UGCCCGUGAGCAUGACCAAGACCAGCGUGGACUGCACCA UGUACAUCUGCGGCGACAGCACCGAGUGCAGCAACCUGC UGCUGCAGUACGGCAGCUUCUGCACCCAGCUGAACCGGG CCCUGACCGGCAUCGCCGUGGAGCAGGACAAGAACACCCA GGAGGUGUUCGCCCAGGUGAAGCAGAUCUACAAGACCCC UCCCAUCAAGGACUUCGGCGGCUUCAACUUCAGCCAGAU CCUGCCCGACCCCAGCAAGCCCAGCAAGCGGAGCUUCAUC GAGGACCUGCUGUUCAACAAGGUGACCCUAGCCGACGCC GGCUUCAUCAAGCAGUACGGCGACUGCCUCGGCGACAUA	

	GCCGCCCGGGACCUGAUCUGCGCCCAGAAGUUCAACGGCC UGACCGUGCUGCCUCCCCUGCUGACCGACGAGAUGAUCGC CCAGUACACCAGCGCCCUGUUAGCCGGAACCAUCACCAGC GGCUGGACUUUCGGCGCUGGAGCCGCUCUGCAGAUCCCC UUCGCCAUGCAGAUGGCCUACCGGUUCAACGGCAUCGGC GUGACCCAGAACGUGCUGUACGAGAACCAGAAGCUGAUC GCCAACCAGUUCAACAGCGCCAUCGGCAAGAUCCAGGAC AGCCUGAGCAGCACCGCUAGCGCCCUGGGCAAGCUGCAG GACGUGGUGAACCAGAACGCCCAGGCCCUGAACACCCUG GUGAAGCAGCUGAGCAGCAACUUCGGCGCCAUCAGCAGC GUGCUGAACGACAUCCUGAGCCGGCUGGACAAGGUGGAG GCCGAGGUGCAGAUCGACCGGCUGAUCACUGGCCGGCUG CAGAGCCUGCAGACCUACGUGACCCAGCAGCUGAUCCGG GCCGCCGAGAUUCGGGCCAGCGCCAACCUGGCCGCCACCA AGAUGAGCGAGUGCGUGCUGGGCCAGAGCAAGCGGGUGG ACUUCUGCGGCAAGGGCUACCACCUGAUGAGCUUUCCCC AGAGCGCACCCCACGGAGUGGUGUUCCUGCACGUGACCU ACGUGCCCGCCCAGGAGAAGAACUUCACCACCGCCCCAGC CAUCUGCCACGACGGCAAGGCCCACUUUCCCCGGGAGGGC GUGUUCGUGAGCAACGGCACCCACUGGUUCGUGACCCAG CGGAACUUCUACGAGCCCCAGAUCAUCACCACCGACAACA CCUUCGUGAGCGGCAACUGCGACGUGGUGAUCGGCAUCG UGAACAACACCGUGUACGAUCCCCUGCAGCCCGAGCUGG ACAGCUUCAAGGAGGAGCUGGACAAGUACUUCAAGAAUC ACACCAGCCCCGACGUGGACCUGGGCGACAUCAGCGGCAU CAACGCCAGCGUGGUGAACAUCCAGAAGGAGAUCGAUCG GCUGAACGAGGUGGCCAAGAACCUGAACGAGAGCCUGAU CGACCUGCAGGAGCUGGGCAAGUACGAGCAGUACAUCAA GUGGCCCUGGUACAUCUGGCUGGGCUUCAUCGCCGGCCU GAUCGCCAUCGUGAUGGUGACCAUCAUGCUGUGCUGCAU GACCAGCUGCUGCAGCUGCCUGAAGGGCUGUUGCAGCUG CGGCAGCUGCUGCAAGUUCGACGAGGACGACAGCGAGCC CGUGCUGAAGGGCGUG	
3′ UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCC CCUUGGGCCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACC CGUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGG C	4
Corresponding amino acid sequence	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVF RSSVLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPF NDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKV CEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVS QPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRD LPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWT AGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKC TLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATR FASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLN DLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTG CVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQ AGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFE LLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKK FLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTN TSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQT RAGCLIGAEHVNNSYECDIPIGAGICASYQTQTVSLRSVASQSII AYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVD CTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQE VFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNK VTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPLLTDE MIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIG	75

	VTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVN QNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLI TGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKR VDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAI CHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFVSG NCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLG DISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIK WPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCLKGCCSCGSCC KFDEDDSEPVLKGV	
PolyA tail	100 nt	
SARS-CoV-2 Variant 32		
SEQ ID NO: 93 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.		93
Chemistry	1-methylpseudouridine	
Cap	7mG(5′)ppp(5′)NlmpNp	
5′ UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUA AGACCCCGGCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGUGAGCAGC CAGUGCGUGAACCUGACCACCCGGACCCAGCUGCCACCAG CCUACACCAACAGCUUCACCCGGGGCGUCUACUACCCCGA CAAGGUGUUCCGGAGCAGCGUCCUGCACAGCACCCAGGA CCUGUUCCUGCCCUUCUUCAGCAACGUGACCUGGUUCCAC GCCAUCCACGUGAGCGGCACCAACGGCACCAAGCGGUUCG ACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUCGC CAGCACCGAGAAGAGCAACAUCAUCCGGGGCUGGAUCUU CGGCACCACCCUGGACAGCAAGACCCAGAGCCUGCUGAUC GUGAAUAACGCCACCAACGUGGUGAUCAAGGUGUGCGAG UUCCAGUUCUGCAACGACCCCUUCCUGGGCGUGUACUACC ACAAGAACAACAAGAGCUGGAUGGAGAGCGAGUUCCGGG UGUACAGCAGCGCCAACAACUGCACCUUCGAGUACGUGA GCCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAGGGCA ACUUCAAGAACCUGCGGGAGUUCGUGUUCAAGAACAUCG ACGGCUACUUCAAGAUCUACAGCAAGCACACCCCAAUCA ACCUGGUGCGGGAUCUGCCCCAGGGCUUCUCAGCCCUGG AGCCCCUGGUGGACCUGCCCAUCGGCAUCAACAUCACCCG GUUCCAGACCCUGCUGGCCCUGCACCGGAGCUACCUGACC CCAGGCGACAGCAGCAGCGGGUGGACAGCAGGCGCGGCU GCUUACUACGUGGGCUACCUGCAGCCCCGGACCUUCCUGC UGAAGUACAACGAGAACGGCACCAUCACCGACGCCGUGG ACUGCGCCCUGGACCCUCUGAGCGAGACCAAGUGCACCCU GAAGAGCUUCACCGUGGAGAAGGGCAUCUACCAGACCAG CAACUUCCGGGUGCAGCCCACCGAGAGCAUCGUGCGGUU CCCCAACAUCACCAACCUGUGCCCCUUCGGCGAGGUGUUC AACGCCACCCGGUUCGCCAGCGUGUACGCCUGGAACCGGA AGCGGAUCAGCAACUGCGUGGCCGACUACAGCGUGCUGU ACAACAGCGCCAGCUUCAGCACCUUCAAGUGCUACGGCG UGAGCCCCACCAAGCUGAACGACCUGUGCUUCACCAACGU GUACGCCGACAGCUUCGUGAUCCGUGGCGACGAGGUGCG GCAGAUCGCACCCGGCCAGACAGGCAAGAUCGCCGACUAC AACUACAAGCUGCCCGACGACUUCACCGGCUGCGUGAUC GCCUGGAACAGCAACAACCUCGACAGCAAGGUGGGCGGC AACUACAACUACCUGUACCGGCUGUUCCGGAAGAGCAAC CUGAAGCCCUUCGAGCGGGACAUCAGCACCGAGAUCUAC CAAGCCGGCUCCACCCCUUGCAACGGCGUGGAGGGCUUCA ACUGCUACUUCCCUCUGCAGAGCUACGGCUUCCAGCCCAC	76

	CAACGGCGUGGGCUACCAGCCCUACCGGGUGGUGGUGCU GAGCUUCGAGCUGCUGCACGCCCCAGCCACCGUGUGUGGC CCCAAGAAGAGCACCAACCUGGUGAAGAACAAGUGCGUG AACUUCAACUUCAACGGCCUUACCGGCACCGGCGUGCUG ACCGAGAGCAACAAGAAAUUCCUGCCCUUUCAGCAGUUC GGCCGGGACAUCGCCGACACCACCGACGCUGUGCGGGAUC CCCAGACCCUGGAGAUCCUGGACAUCACCCCUUGCAGCUU CGGCGGCGUGAGCGUGAUCACCCCAGGCACCAACACCAGC AACCAGGUGGCCGUGCUGUACCAGGACGUGAACUGCACC GAGGUGCCCGUGGCCAUCCACGCCGACCAGCUGACACCCA CCUGGCGGGUCUACAGCACCGGCAGCAACGUGUUCCAGA CCCGGGCCGGUUGCCUGAUCGGCGCCGAGCACGUGAACA ACAGCUACGAGUGCGACAUCCCCAUCGGCGCCGGCAUCUG UGCCAGCUACCAGACCCAGACCGUGUCACUGAGGAGCGU GGCCAGCCAGAGCAUCAUCGCCUACACCAUGAGCCUGGGC GCCGAGAACAGCGUGGCCUACAGCAACAACAGCAUCGCC AUCCCCACCAACUUCACCAUCAGCGUGACCACCGAGAUUC UGCCCGUGAGCAUGACCAAGACCAGCGUGGACUGCACCA UGUACAUCUGCGGCGACAGCACCGAGUGCAGCAACCUGC UGCUGCAGUACGGCAGCUUCUGCACCCAGCUGAACCGGG CCCUGACCGGCAUCGCCGUGGAGCAGGACAAGAACACCCA GGAGGUGUUCGCCCAGGUGAAGCAGAUCUACAAGACCCC UCCCAUCAAGGACUUCGGCGGCUUCAACUUCAGCCAGAU CCUGCCCGACCCCAGCAAGCCCAGCAAGCGGAGCUUCAUC GAGGACCUGCUGUUCAACAAGGUGACCCUAGCCGACGCC GGCUUCAUCAAGCAGUACGGCGACUGCCUCGGCGACAUA GCCGCCCGGGACCUGAUCUGCGCCCAGAAGUUCAACGGCC UGACCGUGCUGCCUCCCCUGCUGACCGACGAGAUGAUCGC CCAGUACACCAGCGCCCUGUUAGCCGGAACCAUCACCAGC GGCUGGACUUUCGGCGCUGGAGCCGCUCUGCAGAUCCCC UUCGCCAUGCAGAUGGCCUACCGGUUCAACGGCAUCGGC GUGACCCAGAACGUGCUGUACGAGAACCAGAAGCUGAUC GCCAACCAGUUCAACAGCGCCAUCGGCAAGAUCCAGGAC AGCCUGAGCAGCACCGCUAGCGCCCUGGGCAAGCUGCAG GACGUGGUGAACCAGAACGCCCAGGCCCUGAACACCCUG GUGAAGCAGCUGAGCAGCAACUUCGGCGCCAUCAGCAGC GUGCUGAACGACAUCCUGAGCCGGCUGGACCCCCCCGAGG CCGAGGUGCAGAUCGACCGGCUGAUCACUGGCCGGCUGC AGAGCCUGCAGACCUACGUGACCCAGCAGCUGAUCCGGG CCGCCGAGAUUCGGGCCAGCGCCAACCUGGCCGCCACCAA GAUGAGCGAGUGCGUGCUGGGCCAGAGCAAGCGGGUGGA CUUCUGCGGCAAGGGCUACCACCUGAUGAGCUUUCCCCA GAGCGCACCCCACGGAGUGGUGUUCCUGCACGUGACCUA CGUGCCCGCCCAGGAGAAGAACUUCACCACCGCCCCAGCC AUCUGCCACGACGGCAAGGCCCACUUUCCCCGGGAGGGCG UGUUCGUGAGCAACGGCACCCACUGGUUCGUGACCCAGC GGAACUUCUACGAGCCCCAGAUCAUCACCACCGACAACAC CUUCGUGAGCGGCAACUGCGACGUGGUGAUCGGCAUCGU GAACAACACCGUGUACGAUCCCCUGCAGCCCGAGCUGGAC AGCUUCAAGGAGGAGCUGGACAAGUACUUCAAGAAUCAC ACCAGCCCCGACGUGGACCUGGGCGACAUCAGCGGCAUCA ACGCCAGCGUGGUGAACAUCCAGAAGGAGAUCGAUCGGC UGAACGAGGUGGCCAAGAACCUGAACGAGAGCCUGAUCG ACCUGCAGGAGCUGGGCAAGUACGAGCAGUACAUCAAGU GGCCCUGGUACAUCUGGCUGGGCUUCAUCGCCGGCCUGA UCGCCAUCGUGAUGGUGACCAUCAUGCUGUGCUGCAUGA CCAGCUGCUGCAGCUGCCUGAAGGGCUGUUGCAGCUGCG GCAGCUGCUGCAAGUUCGACGAGGACGACAGCGAGCCCG	

	UGCUGAAGGGCGUG	
3′ UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCC CCUUGGGCCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACC CGUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGG C	4
Corresponding amino acid sequence	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVF RSSVLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPF NDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKV CEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVS QPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRD LPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWT AGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKC TLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATR FASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLN DLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTG CVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQ AGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFE LLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKK FLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTN TSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQT RAGCLIGAEHVNNSYECDIPIGAGICASYQTQTVSLRSVASQSII AYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVD CTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQE VFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNK VTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPLLTDE MIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIG VTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVN QNAQALNTLVKQLSSNFGAISSVLNDILSRLDPPEAEVQIDRLI TGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKR VDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAI CHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFVSG NCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLG DISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIK WPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCLKGCCSCGSCC KFDEDDSEPVLKGV	77
PolyA tail	100 nt	
SARS-CoV-2 Variant 33		
SEQ ID NO: 94 consists of from 5′ end to 3′ end: 5′ UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 78, and 3′ UTR SEQ ID NO: 4.		94
Chemistry	1-methylpseudouridine	
Cap	7mG(5′)ppp(5′)NlmpNp	
5′ UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUA AGACCCCGGCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGUGAGCAGC CAGUGCGUGAACCUGACCACCCGGACCCAGCUGCCACCAG CCUACACCAACAGCUUCACCCGGGGCGUCUACUACCCCGA CAAGGUGUUCCGGAGCAGCGUCCUGCACAGCACCCAGGA CCUGUUCCUGCCCUUCUUCAGCAACGUGACCUGGUUCCAC GCCAUCCACGUGAGCGGCACCAACGGCACCAAGCGGUUCG ACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUCGC CAGCACCGAGAAGAGCAACAUCAUCCGGGGCUGGAUCUU CGGCACCACCCUGGACAGCAAGACCCAGAGCCUGCUGAUC GUGAAUAACGCCACCAACGUGGUGAUCAAGGUGUGCGAG UUCCAGUUCUGCAACGACCCCUUCCUGGGCGUGUACUACC ACAAGAACAACAAGAGCUGGAUGGAGAGCGAGUUCCGGG	78

	UGUACAGCAGCGCCAACAACUGCACCUUCGAGUACGUGA GCCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAGGGCA ACUUCAAGAACCUGCGGGAGUUCGUGUUCAAGAACAUCG ACGGCUACUUCAAGAUCUACAGCAAGCACACCCCAAUCA ACCUGGUGCGGGAUCUGCCCCAGGGCUUCUCAGCCCUGG AGCCCCUGGUGGACCUGCCCAUCGGCAUCAACAUCACCCG GUUCCAGACCCUGCUGGCCCUGCACCGGAGCUACCUGACC CCAGGCGACAGCAGCAGCGGGUGGACAGCAGGCGCGGCU GCUUACUACGUGGGCUACCUGCAGCCCCGGACCUUCCUGC UGAAGUACAACGAGAACGGCACCAUCACCGACGCCGUGG ACUGCGCCCUGGACCCUCUGAGCGAGACCAAGUGCACCCU GAAGAGCUUCACCGUGGAGAAGGGCAUCUACCAGACCAG CAACUUCCGGGUGCAGCCCACCGAGAGCAUCGUGCGGUU CCCCAACAUCACCAACCUGUGCCCCUUCGGCGAGGUGUUC AACGCCACCCGGUUCGCCAGCGUGUACGCCUGGAACCGGA AGCGGAUCAGCAACUGCGUGGCCGACUACAGCGUGCUGU ACAACAGCGCCAGCUUCAGCACCUUCAAGUGCUACGGCG UGAGCCCCACCAAGCUGAACGACCUGUGCUUCACCAACGU GUACGCCGACAGCUUCGUGAUCCGUGGCGACGAGGUGCG GCAGAUCGCACCCGGCCAGACAGGCAAGAUCGCCGACUAC AACUACAAGCUGCCCGACGACUUCACCGGCUGCGUGAUC GCCUGGAACAGCAACAACCUCGACAGCAAGGUGGGCGGC AACUACAACUACCUGUACCGGCUGUUCCGGAAGAGCAAC CUGAAGCCCUUCGAGCGGGACAUCAGCACCGAGAUCUAC CAAGCCGGCUCCACCCCUUGCAACGGCGUGGAGGGCUUCA ACUGCUACUUCCCUCUGCAGAGCUACGGCUUCCAGCCCAC CAACGGCGUGGGCUACCAGCCCUACCGGGUGGUGGUGCU GAGCUUCGAGCUGCUGCACGCCCCAGCCACCGUGUGUGGC CCCAAGAAGAGCACCAACCUGGUGAAGAACAAGUGCGUG AACUUCAACUUCAACGGCCUUACCGGCACCGGCGUGCUG ACCGAGAGCAACAAGAAAUUCCUGCCCUUUCAGCAGUUC GGCCGGGACAUCGCCGACACCACCGACGCUGUGCGGGAUC CCCAGACCCUGGAGAUCCUGGACAUCACCCCUUGCAGCUU CGGCGGCGUGAGCGUGAUCACCCCAGGCACCAACACCAGC AACCAGGUGGCCGUGCUGUACCAGGACGUGAACUGCACC GAGGUGCCCGUGGCCAUCCACGCCGACCAGCUGACACCCA CCUGGCGGGUCUACAGCACCGGCAGCAACGUGUUCCAGA CCCGGGCCGGUUGCCUGAUCGGCGCCGAGCACGUGAACA ACAGCUACGAGUGCGACAUCCCCAUCGGCGCCGGCAUCUG UGCCAGCUACCAGACCCAGACCAAUUCACCCCGGAGGGCA AGGAGCGUGGCCAGCCAGAGCAUCAUCGCCUACACCAUG AGCCUGGGCGCCGAGAACAGCGUGGCCUACAGCAACAAC AGCAUCGCCAUCCCCCACCAACUUCACCAUCAGCGUGACCA CCGAGAUUCUGCCCGUGAGCAUGACCAAGACCAGCGUGG ACUGCACCAUGUACAUCUGCGGCGACAGCACCGAGUGCA GCAACCUGCUGCUGCAGUACGGCAGCUUCUGCACCCAGCU GAACCGGGCCCUGACCGGCAUCGCCGUGGAGCAGGACAA GAACACCCAGGAGGUGUUCGCCCAGGUGAAGCAGAUCUA CAAGACCCCUCCCAUCAAGGACUUCGGCGGCUUCAACUUC AGCCAGAUCCUGCCCGACCCCAGCAAGCCCAGCAAGCGGA GCUUCAUCGAGGACCUGCUGUUCAACAAGGUGACCCUAG CCGACGCCGGCUUCAUCAAGCAGUACGGCGACUGCCUCGG CGACAUAGCCGCCCGGGACCUGAUCUGCGCCCAGAAGUUC AACGGCCUGACCGUGCUGCCUCCCCUGCUGACCGACGAGA UGAUCGCCCAGUACACCAGCGCCCUGUUAGCCGGAACCAU CACCAGCGGCUGGACUUUCGGCGCUGGAGCCGCUCUGCA GAUCCCCUUCGCCAUGCAGAUGGCCUACCGGUUCAACGGC AUCGGCGUGACCCAGAACGUGCUGUACGAGAACCAGAAG	

	CUGAUCGCCAACCAGUUCAACAGCGCCAUCGGCAAGAUCC AGGACAGCCUGAGCAGCACCGCUAGCGCCCUGGGCAAGC UGCAGGACGUGGUGAACCAGAACGCCCAGGCCCUGAACA CCCUGGUGAAGCAGCUGAGCAGCAACUUCGGCGCCAUCA GCAGCGUGCUGAACGACAUCCUGAGCCGGCUGGACCCUCC CGAGGCCGAGGUGCAGAUCGACCGGCUGAUCACUGGCCG GCUGCAGAGCCUGCAGACCUACGUGACCCAGCAGCUGAU CCGGGCCGCCGAGAUUCGGGCCAGCGCCAACCUGGCCGCC ACCAAGAUGAGCGAGUGCGUGCUGGGCCAGAGCAAGCGG GUGGACUUCUGCGGCAAGGGCUACCACCUGAUGAGCUUU CCCCAGAGCGCACCCCACGGAGUGGUGUUCCUGCACGUGA CCUACGUGCCCGCCCAGGAGAAGAACUUCACCACCGCCCC AGCCAUCUGCCACGACGGCAAGGCCCACUUUCCCCGGGAG GGCGUGUUCGUGAGCAACGGCACCCACUGGUUCGUGACC CAGCGGAACUUCUACGAGCCCCAGAUCAUCACCACCGACA ACACCUUCGUGAGCGGCAACUGCGACGUGGUGAUCGGCA UCGUGAACAACACCGUGUACGAUCCCCUGCAGCCCGAGCU GGACAGCUUCAAGGAGGAGCUGGACAAGUACUUCAAGAA UCACACCAGCCCCGACGUGGACCUGGGCGACAUCAGCGGC AUCAACGCCAGCGUGGUGAACAUCCAGAAGGAGAUCGAU CGGCUGAACGAGGUGGCCAAGAACCUGAACGAGAGCCUG AUCGACCUGCAGGAGCUGGGCAAGUACGAGCAGUACAUC AAGUGGCCCUGGUACAUCUGGCUGGGCUUCAUCGCCGGC CUGAUCGCCAUCGUGAUGGUGACCAUCAUGCUGUGCUGC AUGACCAGCUGCUGCAGCUGCCUGAAGGGCUGUUGCAGC UGCGGCAGCUGCUGCAAGUUCGACGAGGACGACAGCGAG CCCGUGCUGAAGGGCGUG	
3′ UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCC CCUUGGGCCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACC CGUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGG C	4
Corresponding amino acid sequence	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVF RSSVLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPF NDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKV CEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVS QPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRD LPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWT AGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKC TLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATR FASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLN DLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTG CVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQ AGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFE LLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKK FLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTN TSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQT RAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPRRARSVA SQSIIAYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTK TSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDK NTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDL LFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPL LTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRF NGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQ DVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDPPEAEVQ IDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLG QSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTT APAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNT FVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPD	79

	VDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYE QYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCLKGCCSC GSCCKFDEDDSEPVLKGV	
PolyA tail	100 nt	
CoV2 Membrane Protein		
SEQ ID NO: 95 consists of from 5′ end to 3′ end: 5′ UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 80, and 3′ UTR SEQ ID NO: 4.		95
Chemistry	1-methylpseudouridine	
Cap	7mG(5′)ppp(5′)NlmpNp	
5′ UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUA AGACCCCGGCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGGCCGACAGCAACGGCACCAUCACCGUGGAGGAGCUG AAGAAGCUGCUGGAGCAGUGGAACCUGGUGAUCGGCUUC CUGUUCCUGACCUGGAUCUGCCUGCUGCAGUUCGCCUAC GCCAACCGGAACCGUUUCCUGUACAUCAUCAAGCUGAUC UUCCUGUGGCUGCUGUGGCCCGUGACCCUGGCCUGCUUC GUGCUGGCCGCCGUGUACCGGAUCAACUGGAUCACCGGC GGCAUCGCCAUCGCCAUGGCCUGCCUGGUGGGCCUGAUG UGGCUGAGCUACUUCAUCGCCAGCUUCCGGCUGUUCGCCC GGACCCGGAGCAUGUGGAGCUUCAACCCCGAGACCAACA UCCUGCUGAACGUGCCCCUGCACGGCACAAUCCUGACCCG GCCCCUGCUGGAGAGCGAGCUUGUGAUCGGCGCCGUGAU CCUGCGGGGCCACCUGAGGAUCGCCGGCCAUCACCUGGGC CGGUGCGACAUCAAGGACCUGCCCAAGGAGAUCACCGUG GCCACCAGCCGGACCCUGAGCUACUACAAACUGGGCGCCA GCCAGAGAGUGGCCGGAGACAGCGGCUUCGCCGCCUACA GCCGGUACCGGAUCGGCAACUACAAGCUGAACACCGACC ACAGCAGCAGCAGCGACAACAUCGCCCUGCUGGUGCAG	80
3′ UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCC CCUUGGGCCUCCCCCCCAGCCCCUCCUCCCCUUCCUGCACC CGUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGG C	4
Corresponding amino acid sequence	MADSNGTITVEELKKLLEQWNLVIGFLFLTWICLLQFAYANR NRFLYIIKLIFLWLLWPVTLACFVLAAVYRINWITGGIAIAMAC LVGLMWLSYFIASFRLFARTRSMWSFNPETNILLNVPLHGTIL TRPLLESELVIGAVILRGHLRIAGHHLGRCDIKDLPKEITVATS RTLSYYKLGASQRVAGDSGFAAYSRYRIGNYKLNTDHSSSSD NIALLVQ	81
PolyA tail	100 nt	
CoV2 Envelope Protein		
SEQ ID NO: 96 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.		96
Chemistry	1-methylpseudouridine	
Cap	7mG(5′)ppp(5′)NlmpNp	
5′ UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUA AGACCCCGGCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGUACAGCUUCGUGAGCGAGGAGACCGGCACCCUGAUC GUGAACAGCGUGCUGCUGUUCCUGGCCUUCGUGGUGUUC CUGCUGGUGACCCUGGCCAUCCUGACCGCCCUGCGGCUGU GUGCCUACUGCUGCAACAUCGUGAACGUGAGCCUGGUGA AGCCCAGCUUCUACGUGUACAGCCGGGUGAAGAACCUGA ACAGCAGCCGGGUGCCUGACCUGCUGGUG	82

3′ UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCC CCUUGGGCCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACC CGUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGG C	4
Corresponding amino acid sequence	MYSFVSEETGTLIVNSVLLFLAFVVFLLVTLAILTALRLCAYCC NIVNVSLVKPSFYVYSRVKNLNSSRVPDLLV	83
PolyA tail	100 nt	
CoV2 Nucleocapsid Protein		
SEQ ID NO: 97 consists of from 5′ end to 3′ end: 5′ UTR SEQ ID NO: 2, mRNA ORF SEQ ID NO: 84, and 3′ UTR SEQ ID NO: 4.		97
Chemistry	1-methylpseudouridine	
Cap	7mG(5′)ppp(5′)NlmpNp	
5′ UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUA AGACCCCGGCGCCGCCACC	2
ORF of mRNA Construct (excluding the stop codon)	AUGAGCGACAACGGCCCUCAGAACCAGCGGAACGCACCCC GGAUCACCUUUGGCGGCCCCAGCGAUAGCACCGGCAGCA ACCAGAACGGCGAGAGAUCAGGGGCCCGGAGCAAGCAGC GGCGUCCUCAGGGCCUGCCCAACAACACCGCCAGCUGGUU CACCGCCCUGACCCAGCACGGCAAGGAGGACCUGAAGUUC CCUCGGGGCCAAGGAGUGCCCAUCAACACCAACAGCAGCC CCGACGACCAGAUCGGCUACUACAGAAGGGCCACCCGGA GGAUCCGGGGAGGGGACGGCAAGAUGAAGGACCUGUCUC CCCGGUGGUACUUCUACUAUCUUGGCACGGGCCCUGAAG CUGGCCUGCCGUACGGCGCAAACAAGGACGGCAUCAUCU GGGUCGCCACCGAGGGAGCCCUGAACACCCCGAAGGACCA CAUCGGCACCCGGAAUCCCGCCAACAACGCCGCCAUCGUU CUGCAGCUGCCCCAGGGCACCACCCUGCCCAAGGGCUUCU ACGCCGAGGGCAGCAGAGGCGGCUCACAGGCCAGCAGCC GGUCAAGCAGCCGGAGCCGGAACAGCAGCCGGAACUCCA CACCCGGCUCUAGCCGAGGCACAAGCCCCGCCAGAAUGGC AGGAAACGGCGGCGACGCUGCCUUAGCCCUGCUGUUGCU GGACCGGCUGAACCAGCUCGAGAGCAAGAUGAGCGGCAA GGGUCAGCAGCAGCAAGGCCAAACCGUGACCAAGAAGAG CGCCGCCGAGGCUAGCAAGAAGCCCCGGCAGAAGCGGACC GCCACCAAGGCCUACAACGUGACCCAGGCCUUCGGUCGGA GAGGCCCCGAGCAGACCCAGGGCAACUUCGGCGACCAGG AGCUGAUCCGGCAGGGCACCGACUACAAGCACUGGCCCCA GAUCGCCCAGUUCGCCCCUAGCGCCUCAGCCUUCUUCGGC AUGAGCCGGAUCGGCAUGGAGGUGACUCCCAGCGGCACC UGGCUGACCUACACCGGCGCCAUCAAGCUGGACGACAAG GACCCCAACUUCAAGGACCAGGUGAUCCUGCUGAACAAG CACAUCGACGCCUACAAGACCUUUCCGCCCACCGAGCCCA AGAAGGACAAGAAGAAGAAGGCCGACGAGACCCAGGCCC UGCCCCAACGGCAGAAGAAGCAGCAGACCGUCACCUUAC UGCCCGCAGCCGACCUGGACGACUUCAGCAAGCAGCUGCA GCAGAGCAUGAGCAGCGCCGACAGCACCCAGGCC	84
3′ UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCC CCUUGGGCCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACC CGUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGG C	4
Corresponding amino acid sequence	MSDNGPQNQRNAPRITFGGPSDSTGSNQNGERSGARSKQRRP QGLPNNTASWFTALTQHGKEDLKFPRGQGVPINTNSSPDDQIG YYRRATRRIRGGDGKMKDLSPRWYFYYLGTGPEAGLPYGAN KDGIIWVATEGALNTPKDHIGTRNPANNAAIVLQLPQGTTLPK GFYAEGSRGGSQASSRSSSRSRNSSRNSTPGSSRGTSPARMAG	85

	NGGDAALALLLLDRLNQLESKMSGKGQQQQGQTVTKKSAAE ASKKPRQKRTATKAYNVTQAFGRRGPEQTQGNFGDQELIRQG TDYKHWPQIAQFAPSASAFFGMSRIGMEVTPSGTWLTYTGAI KLDDKDPNFKDQVILLNKHIDAYKTFPPTEPKKDKKKKADET QALPQRQKKQQTVTLLPAADLDDFSKQLQQSMSSADSTQA	
PolyA tail	100 nt	

* It should be understood that any one of the open reading frames and/or corresponding amino acid sequences described in Table 1 may include or exclude the signal sequence. It should also be understood that the signal sequence may be replaced by a different signal sequence, for example, any one of SEQ ID NOs: 38-43.

EQUIVALENTS

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.