(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
12 August 2021 (12.08.2021)

WIPO | PCT

(10) International Publication Number
WO 2021/159118 A3

(51) International Patent Classification:

A61K 48/00 (2006.01)	*A61K 39/12* (2006.01)
A61K 9/127 (2006.01)	*A61K 39/00* (2006.01)
A61K 9/00 (2006.01)	*A61K 38/00* (2006.01)
A61K 9/14 (2006.01)	*C07H 21/04* (2006.01)

(21) International Application Number:
PCT/US2021/017149

(22) International Filing Date:
08 February 2021 (08.02.2021)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

62/971,834	07 February 2020 (07.02.2020)	US
63/058,463	29 July 2020 (29.07.2020)	US
63/130,581	24 December 2020 (24.12.2020)	US

(71) Applicant: **RNALMMUNE, INC.** [US/US]; 401 Professional Drive, Suite 280, Gaithersburg, MD 20879 (US).

(72) Inventors: **TANG, Shenggao**; c/o Rnalmmune, Inc., 401 Professional Drive, Suite 280, Gaithersburg, MD 20879 (US). **SHEN, Dong**; c/o Rnalmmune, Inc., 401 Professional Drive, Suite 280, Gaithersburg, MD 20879 (US). **LU, Chun**; c/o Rnalmmune, Inc., 401 Professional Drive, Suite 280, Gaithersburg, MD 20879 (US). **HE, Ziyang**; c/o Rnalmmune, Inc., 401 Professional Drive, Suite 280, Gaithersburg, MD 20879 (US). **HE, Jiaxi**; c/o Rnalmmune, Inc., 401 Professional Drive, Suite 280, Gaithersburg, MD 20879 (US). **LU, Patrick, Y.**; c/o Rnalmmune, Inc., 401 Professional Drive, Suite 280, Gaithersburg, MD 20879 (US).

(74) Agent: **BOOTH, Paul, M.** et al.; Edell, Shapiro & Finnan LLC, 9801 Washingtonian Blvd., Suite 750, Gaithersburg, MD 20878 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report (Art. 21(3))*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

(88) Date of publication of the international search report:
28 October 2021 (28.10.2021)

(54) Title: COMPOSITION AND METHOD OF MRNA VACCINES AGAINST NOVEL CORONAVIRUS INFECTION

(57) Abstract: Compositions and methods are provided for potent mRNA vaccines for prevention and treatment of 2019 novel Coronavirus (2019-nCoV) infections. The compositions include a pharmaceutical composition containing one or more mRNA molecules encoding spike protein epitopes, including mutated epitopes, or mRNA cocktails that encode critical viral genes together with pharmaceutically acceptable polymeric nanoparticle carriers and liposomal nanoparticle carriers. Methods for stimulating system immune responses and treatment are provided, including subcutaneous, intraperitoneal and intramuscular injections.

WO 2021/159118 A3

COMPOSITION AND METHODS OF mRNA VACCINES AGAINST NOVEL CORONAVIRUS INFECTION

This application claims priority to U.S.Provisional application nos. 62/971,834, filed February 7, 2020, 63/058,463, filed July 29, 2020, and 63/130,581, filed December 24, 2020, the contents of each of which are hereby incorporated by reference in their entireties.

FIELD OF INVENTION

Prophylactic and therapeutic agents for vaccination, prevention and treatment of 2019-nCoV infection are provided, together with methods for their use.

BACKGROUND

Novel coronavirus (2019-nCoV) infection: Biology and Pathology

Coronaviruses (CoVs) have repeatedly crossed species barriers and some have emerged as important human pathogens. During the past two decades, two coronaviruses infecting animals have evolved and caused outbreaks in humans: SARS-CoV (2002, Betacoronavirus, subgenus Sarbecovirus), and MERS-CoV (2012, Betacoronavirus, subgenus Merbecovirus) [Drosten *et al., New Engl J Med.* 2003; 348: 1967-1976; Zaki *et al.,. New Engl J Med.* 2012; 367: 1814-1820]. On 31 December 2019, the Wuhan Municipal Health Commission in Wuhan City, Hubei province, China reported a cluster of 27 pneumonia cases of unknown etiology, including seven severe cases, with a common reported link to Wuhan's Huanan Seafood Wholesale Market (a wholesale fish and live animal market selling different animal species) [Wuhan City Health Committee (WCHC). Wuhan Municipal Health and Health Commission's briefing on the current pneumonia epidemic situation in our city 2019 (updated 31 December 201914 January 2020). Available from: http://wjw.wuhan.gov.cn/front/web/showDetail/2019123108989]. The 2019-nCoV is a new strain of coronavirus that has not been previously identified in humans [Zhu *et al., N Engl J Med.* 382:727-733 (2020)]. The outbreak has rapidly spread to affect other parts of China and outside the country. Cases have now been detected in several countries in Asia, but also in Australia, Europe and North America. It appears that the 2019-nCoV can be spread through

human-human contact via respiratory droplets [Li *et al.*, *N Engl J Med.* 2020; 382:1199-1207]..

Genetic analysis revealed that 2019-nCoV is closely related to SARS-CoV and genetically clusters within the genus Betacoronavirus, forming a distinct clade in lineage B of the subgenus Sarbecovirus together with two bat-derived SARS-like strains. The virus genome consists of six major open reading frames (ORFs) common to coronaviruses and a number of other accessory genes. Further analysis indicates that some of the 2019-nCoV genes shared less than 80% sequence identity to SARS-CoV. However, the seven conserved replicase (RdRP) domains in ORF1ab that were used for CoV species classification, are 94.6% AA sequence identical between 2019-nCoV and SARS-CoV, implying the two belong to the same species. Angiotensin converting enzyme II (ACE2), a membrane exopeptidase, was known as the cell receptor for SARS- CoV. This receptor also has been shown to be used by 2019-nCoV for entry into human cells. Domain 266–330 (numbering based on the sequence of the mature protein) has been shown to be essential for interaction of 2019-nCoV with ACE2 [Veljkovic *et al.*, F1000Research. 2020; 9: 52.].

Current Prophylaxis and Therapeutics for Novel coronavirus (2019-nCoV)

At present, the sequences of 287,543 full genomes of 2019-nCoV have been submitted on GISAID [GISAID, Newly discovered betacoronavirus, Wuhan 2019-2020 (2019); https://www.gisaid.org/.] by China CDC, *etc.*, including one that has been released on GenBank(Accession: MN908947.3) [http://www.ncbi.nlm.nih.gov/genbank/.]. Currently, there are only two emergency use authorized vaccines and no licensed vaccines for preventing coronavirus infections, including 2019-nCoV infections.

Advantages of mRNA vaccines

Vaccination is the most successful medical approach to disease prevention and control. The successful development and use of vaccines has saved thousands of lives and large amounts of money. A key advantage of RNA vaccines is that RNA can be produced in the laboratory from a DNA template using readily available materials, less expensively and

faster than conventional vaccine production, which can require the use of chicken eggs or other mammalian cells. In addition, mRNA vaccines have the potential to streamline vaccine discovery and development, and facilitate a rapid response to emerging infectious diseases [Maruggi *et al., Mol Ther.* 2019; 27(4): 757-772

Preclinical and clinical trials have shown that mRNA vaccines provide a safe and long-lasting immune response in animal models and humans. mRNA vaccines against infectious diseases may be developed as prophylactic or therapeutic treatments. mRNA vaccines expressing antigens of infectious pathogens have been shown to induce potent T cell and humoral immune responses [Pardi *et al.,Nat Rev Drug Discov.* 2018; 17: 261–279.]. The production procedure to generate mRNA vaccines is cell-free, simple, and rapid, compared to production of whole microbe, live attenuated, and subunit vaccines. This fast and simple manufacturing process makes mRNA a promising bio-product that can potentially fill the gap between emerging infectious disease and the desperate need for effective vaccines.

SUMMARY OF THE INVENTION

Compositions are provided containing an effective amount of a messenger ribonucleic acid (mRNA) that contains an open reading frame (ORF) encoding a 2019-nCoV protein or protein fragment formulated in a pharmaceutically acceptable carrier, where the mRNA and the carrier form a polymeric nanoparticle. The 2019-nCoV protein may be a spike (S) protein or spike (S) subunit. The ORF may be an mRNA encoding a peptide such as the peptides having a sequence of selected from SEQ ID 14 to 31, or 50-66. The ORF may contain for example, an mRNA have a sequences as shown in SEQ ID Nos. 32 to 49 or SEQ ID No.91 to 107. The ORF may be selected from mRNA molecules encoding the amino acid sequence selected from SEQ ID No. 11, 13, 68, 70, 72, 74 and 90. The ORF may have a sequence such as SEQ ID No. 1-4, 10, 12, 67, 69, 71, 73 or 89.

These composition may also contain an mRNA containing all or part of orf1a and/or orf1b, where the mRNA encodes a 2019-nCoV non-structural protein polymerase and/or RdRp. The ORF may be an mRNA having a sequence such as SEQ ID NO:. 5 or 6. The

ORF may encode all or part of the 2019 nCoV encoding envelope (E), membrane (M), and/or nucleocapsid (N) protein. These compositions may also contain one or more mRNA sequences encoding all or part of the 2019 nCoV encoding envelope (E), membrane (M), and/or nucleocapsid (N) protein. The ORF may be selected from mRNA molecules having a sequence shown in SEQ ID No.7, 8, or 9.

In any of these compositions, the mRNA may be chemically modified with one or more N1-methyl-pseudouridine and/or pseudouridine residues. The uridine residues in the mRNA may all be N1-methyl pseudouridine or pseudouridine.

In any of these compositions the mRNA molecule may further contain a 3'-UTR, a 5'-UTR and optionally further contains additional elements that (a) stabilize the composition and (b) enhance expression of the polypeptide encoded by the ORF. The 5'-UTR may contain an m7G cap structure and a start codon and/or the 3'-UTR may contain a stop codon and a polyA tail.

The composition may contain more than one ORF and may encode at least two polypeptides.

The composition may contain a polymeric nanoparticle carrier that contains a Histidine-Lysine co-polymer (HKP). Examples of suitable HKP include, but are not limited to polymers having side chains selected from SEQ ID NO:75-88.

Also provided are methods of making a composition as described above by expressing the ORF in a linear *in vitro* transcription (IVT) system or by a plasmid DNA (pDNA) vector delivery system.

The compositions as described above may be made by contacting a composition containing the ORF(s) with an HKP, where the mRNA and the HKP self-assemble into nanoparticles. The composition may further contain a lipid, where the lipid optionally may be a cationic lipid.

Also provided are methods of treating a subject suffering from a 2019-nCoV infection by administering to the subject a pharmaceutically effective amount of a composition as described above. The composition may be atomized by a nebulizer inhalation system prior

to or during administration. The nebulizer system may be a portable nebulizer for whole respiratory tract drug delivery. The composition may be administered by subcutaneous injection, intramuscular injection, or intraperitoneal injection (i.p).

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows a schematic representation of optimized mRNA vaccine expression structure. This structure is contained in a linear in vitro transcription (IVT) expression system or plasmid DNA delivery vector.

Figure 2 shows immunofluorescence analysis result of spike protein expression after mRNA transient transfection.

Figure 3 shows Western Blot result of Spike Protein expression after mRNA transient transfection.

Figure 4 shows spike protein ORF mRNA sequences

Figure 5 shows D614G mutated spike protein ORF mRNA and protein sequences

Figure 6 shows SARS RBD mRNA sequence, and the amino acid sequence for D614G Mutated Spike Protein ORF

Figure 7 shows mRNA and amino acid sequences for polypeptide epitopes for 2019-nCoV

Figure 8 shows the D614G_S6P _DNA and amino acid sequences

Figure 9 shows RBD trimer DNA and amino acid sequences

Figure 10 shows VUI-202012 DNA and amino acid sequences

Figure 11 shows HKP peptide sequences

Figure 12 shows D614G S2P DNA Sequence

Figure 13 shows D614G S2P amino acid sequence

Figure 14 shows the sequences of mRNA molecules encoding peptide epitopes of 2019-nCoV.

DETAILED DESCRIPTION

Prophylactic and therapeutic agents for vaccination, prevention and treatment of 2019-nCoV infection are provided. Pharmaceutical compositions are provided that contain mRNA sequences formulated with pharmaceutical excipients. The composition of the mRNA-based pharmaceutical formulation contains mRNA sequences coding viral proteins necessary for coronavirus infection and replication, and a specific carrier. Suitable carriers include Histidine-Lysine Co-polymers (HKP), and/or HKP plus a lipid, such as a cationic lipid. Methods also are provided for using these pharmaceutical compositions, including methods of treatment, process development, and specific delivery routes and regimens.

Target selection and design for mRNA vaccine against 2019-nCoV

In regard to the protein level, the predicted genes of 2019-nCoVs were identified and annotated. It was found that the annotation results of all six 2019-nCoV genomes are identical, including seven functional proteins (pp1a, E2 glycoprotein precursor, hypothetical protein sars3a, matrix protein, hypothetical protein sars6, hypothetical protein sars7a, nucleocapsid protein) verified in SARS-CoV-like coronavirus [B. Buchfink, *et al., Nat Methods*. 2015; 12: 59-60]. These results illustrate the similarities and differences between 2019-nCoV and SARS-CoV or MERS-CoV-like Coronaviruses.

The predominant human receptor for 2019-nCoV is believed to be human angiotensin-converting enzyme 2 (ACE2). Domain 288-330 of S1 protein from the 2019-nCoV represents a promising therapeutic and/or vaccine target [Veljkovic *et al., supra*]. Generally, the spike (S) protein of coronavirus including 2019-nCoV is the major inducer of neutralizing antibodies, and the receptor-binding domain (RBD) in the S1 subunit of S protein contains multiple conformational neutralizing epitopes [He *et al., J Immunol.* 2005; 174(8): 4908-4915]. This suggests that recombinant proteins containing RBD and vectors including mRNAs encoding the RBD sequence can be used to develop safe and effective 2019-nCoV vaccines.

As described herein, a new type of mRNA vaccine (mRNA-1273) has been developed using an immunogenic composition that contains a messenger ribonucleic acid (mRNA)

containing an open reading frame (ORF) encoding one or more epitopes from one or more proteins of 2019-nCoV. The mRNA is formulated in a pharmaceutically acceptable carrier containing a polymeric nanoparticle and/or a liposomal nanoparticle. The composition may be administered to a subject in an amount effective to induce a specific immune response against 2019-nCoV in the subject. The composition may contain more than one type of mRNA molecule. Each additionaly copy enhances the immunostimulatory activities of the composition. All the mRNAs encode antigens that are highly immunogenic and that are required for 2019-nCoV to enter epithelial cells, which is the first step in 2019-nCoV infection.

In one embodiment, the composition contains one or more mRNAs encoding the entire spike (S) protein of 2019-nCoV (Fig. 3). The S protein is used by the virus to attach to the host cell receptor. The Spike protein-coding mRNA sequence is long, and therefore in order to improve the efficiencies of mRNA expression systems preparation, it was divided into four regions, where one, some, or all of these regions can be used as mRNA antigen(s) against Spike protein structure (SEQ ID Nos.1 to 4), when expressed in frame for the Spike protein domains. 24 short peptide fragments that are potential epitopes were selected as targets (SEQ ID Nos. 14 to 31). The short peptide selected is less than 20 amino acids, advantageously less than 15 amino acids. mRNA sequences encoding short peptide sequences were designed using methods that are well known in the art (SEQ ID No.32 to 49).

Furthermore, compared to other coronavirus, 2019-nCoV has several unique insertions in the spike protein. The amino acid residues in those insertion sites of S protein of 2019-nCoV are identical or similar to the amino acid residues in the replicating protein gp120 or Gag of human immunodeficiency virus type 1 (HIV-1). In another embodiment, a series of mRNA sequences was designed that encode amino acid residues that are identical between 2019-nCoV and HIV-1 (SEQ ID No.50 to 66), which are recognized by the human immune system.

In another embodiment, the mRNA encodes the D614G variant of the ORF of the spike protein. The coronavirus disease 2019 (COVID-19) outbreak, caused by SARS-CoV-2,

has rapidly expanded to a global pandemic. SARS-CoV-2 isolates encoding a D614G mutation in the viral spike (S) protein predominate over time in locales where it is found, implying that this change enhances viral transmission. S proteins with aspartic acid (D614) and glycine (G614) at residue 614 have different functional properties. Specifically, retroviruses pseudotyped with G614 were found to infect ACE2- expressing cells much more efficiently than those with wild type (D614). This greater infectivity was correlated with less S1 shedding and greater incorporation of the S protein into the pseudovirion. Similar results were obtained using virus-like particles produced with SARS-CoV-2 M, N, E, and S proteins. In correlation analyses, spike protein bearing the D614G mutation showed significant positive correlations with fatality rates ($r = 0.43$, $P = 0.022$). These results indicate that spike protein with the D614G mutation is more stable than wild type, which is consistent with epidemiological data suggesting that viruses with the spike D614G mutation transmit more efficiently (SEQ ID No. 10).A vaccine with an ORF that incorporates the D614G mutation was produced because it would more effectively protect against the dominant strain of Sars-Cov-2.

In another embodiment, the mRNA encodes the ORF with SEQ ID NO:89, encoding D614G spike protein precursor S2P.

In another embodiment, the mRNA encodes the ORF with Seq ID No.67 , which incorporates other mutations in the spike protein that have been identified. These include the HV 69-70 deleted spike protein, the Y144 deleted spike protein, the N501Y mutated spike protein, the A570D mutated spike protein, the P681H mutated spike protein, the T716I mutated spike protein, the F817P modified spike protein, the A892P modified spike protein, the A899P modified spike protein, the A942P modified spike protein, the S982A mutated spike protein, and the D1118H mutated spike protein. Vaccine compositions containing mRNAs encoding one or more of any of these known mutations can be prepared using the methods described herein. .

In other embodiments, the mRNA may encode one or more other SARS-CoV-2 viral protein coding sequences including the Envelope (E), Membrane (M), or Nucleocapsid (N)

structural proteins. The 2019- nCoV genome is arranged in the order of: 5′ -replicase (orf1/ab)-structural proteins [Spike (S)-Envelope (E)-Membrane (M)-Nucleocapsid (N)]−3′ (Zulma et al. 2016, Nature Reviews Drug Discovery 15:327-347.), and lacks the hemagglutinin-esterase gene which is characteristically found in lineage A β-CoV. Therefore, various viral proteins other than, or in addition to, the S protein of 2019-nCoV coding mRNA can be applied as the mRNA based vaccination targets. In an embodiment, the viral surface proteins advantageously are targets for an mRNA vaccine, The viral functional proteins responsible for viral polymerase in the region of ORF1a (SEQ ID No. 5) and replication (RdRp) in the region of ORF1b (SEQ ID No. 6) also may be used to prepare a vaccine composition.

Other viral proteins may also be the target of mRNA vaccines. For example, the coronavirus membrane (M) protein is a key player in virion assembly. One of its functions is to mediate the incorporation of the spikes into the viral envelope. The nucleocapsid N-protein is a highly immunogenic phosphoprotein, that is typically highly conserved. The coronavirus N-protein is often used as a marker in diagnostic assays. Coronavirus (CoV) envelope (E) proteins play multiple roles during infection, including virus morphogenesis and pathogenesis. In other embodiments, mRNAs coding for all or part of one or more viral M protein, N protein and E protein as their specific antigens may be selected for use as components in the antigenic vaccine composition (SEQ ID No.7,8,9).

Modifications may be made in the mRNA sequence to prepare the vaccine composition. For example, a 5’ cap structure may be added to the mRNA sequence. In one embodiment, a 5’CleanCap is added. This structure uses an initiating capped trimer to yield a naturally occurring 5’ cap structure. Advantageously,a poly-A tail is added to the sequence to increase mRNA stability. AU or GU-enriched sequences may be selected in the 3’ UTR, while random sequences may be utilized in the 5’ UTP to reduce similarity to the host gene sequences. Codon optimization of the mRNA sequence may also be used to improve translation efficiency. Specifically, codons that correspond to lower abundance tRNA species and that therefore lead to lower protein expression are replaced with codons that correspond

to higher abundance tRNAs that provide higher levels of protein expression. (Brown, *et al. Protein Expr Purif*, 2008, 59:94–102; Maertens *et al.*, 2010, *Protein Sci,* 9:1312–1326.; Plotkin *et al.*, 2011, *Nat Rev Genet,* 12:32–42.; Kudla *et al.*, 2009, *Science,* 324:255–258.)

Improving mRNA Vaccine Expression efficiency

To improve mRNA expression efficiency in mammalian cells, mRNA stability may be enhanced by partial chemical modification. To further increase the translation efficiency, short and double strand RNAs derived from aberrant RNA polymerase activities are removed, for example by FPLC, HPLC, or chromatography using cellulose as a matrix. To improve the potency of mRNA vaccines, sequence optimization may be used, together with usage of modified nucleosides, such as pseudouridine (φ), 5-methylcytidine (5mC), Cap-1 structure and optimized codons, which in turn improve translation efficiency. [Do we have any more information about the use of modified nucleosides? Or is that common knowledge? I think codon optimization is common knowledge at this point] During *in vitro* transcription of mRNA, immature mRNA may be produced as a contaminant which inhibits translation through stimulating innate immune activation (Kariko et al, *Mol Ther*. 2008 16(11):1833-40.; Andries et al, *J Control Release*. 2015 Nov 10;217:337-44.). FPLC and HPLC purification may be used to remove these contaminants.

In the compositions described herein, the template for *in vitro* transcription of mRNA contains five cis-acting structural elements, namely from 5' to 3': (i) the optimized cap structure, (ii) the optimized 5' untranslated region (UTR), (iii) the codon optimized coding sequence, (iv) the optimized 3' UTR and (v) a stretch of repeated adenine nucleotides (polyA tail) (Fig. 18). (Murray et al, Mol. *Cell. Biol.* 2007, 27, 2791–2799. ; Louis et al, *Curr. Opin. Genet. Dev.* 2011, 21, 444–451. ; Ferizi et al, *Lab Chip* 2015, 15, 3561–3571. ; Von Niessen et al, *Mol. Ther*. 2019, 27, 824–836.) These cis-acting structural elements are further optimized in the endeavor for better mRNA features, including half-life and expression level. Provided herein, the 5'-UTR includes a start, but does not encode polypeptide (*i.e.* it is non-coding). In some embodiments, a 5'-UTR of the present disclosure comprises a cap structure with 7-methylguanosine (7mG) sequences. The 3'-UTR is directly downstream (3') from the

stop codon (the codon of an mRNA transcript representing a termination signal) and does not encode a polypeptide (is non-coding). A polyA tail is a special region of mRNA that is downstream from 3'-UTR and contains multiple consecutive adenosine monophosphates; it is required for efficient mRNA production. An mRNA *in vitro* transcription template is used not only for efficiency of mRNA production but also for the subsequent protein or peptide production

In some embodiments, mRNA is produced by *in vitro* transcription (IVT) from a linear DNA template containing a bacteriophage promoter, the optimized UTR's and the codon optimized sequence by using a RNA polymerase (T7, T3 or SP6) and a mix of the different nucleosides. In other embodiments, the linear DNA template can be cloned into a plasmid DNA (pDNA) as a delivery vector. The plasmid vectors can be adapted for mRNA vaccine production. Commonly used plasmids include pSFV1, pcDNA3 and pTK126, which are all commercially available. One unique mRNA expression system is pEVL (see Grier *et al. Mol Ther Nucleic Acids*. 19;5:e306 (2016),

In some embodiments, the mRNA is a modified mRNA where all uridine residues are replaced by pseudouridines during *in vitro* transcription. This modification stabilizes the mRNA against enzymatic degradation in the cell, leading to enhanced translation efficiency of the mRNA. The pseudouridines used can be N1-methyl-pseudouridine, or other modifications that are well known in the art such as N6 -methyladenosine (m6A), inosine, pseudouridine, 5-methylcytidine (m5C), 5-hydroxymethylcytidine (hm5C), and N1-methyladenosine (m1A). The modification optionally is made throughout the entire mRNA, including the ORF, the 5'UTR and 3'UTR. The skilled artisan will recognize that other modified RNA residues may be used to stabilize the the protein 3 dimensional structure and increase protein translation.

To test the level of efficiency, the experiments described in Example 3, 4 and 5 were performed.

Histidine-Lysine (HK) polypeptides as mRNA vaccine delivery systems

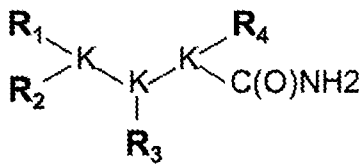
Effective means for transferring nucleic acids into target cells are important tools, both in the basic research setting and in clinical applications. A diverse array of nucleic acid carriers is currently required because the effectiveness of a particular carrier depends on the characteristics of the nucleic acids that is being transfected [Blakney et al. *Biomacromolecules* 2018, **19**: 2870-2879. Goncalves et al. *Mol Pharm* 2016; **13**: 3153-3163. Kauffman et al. *Biomacromolecules* 2018; **19**: 3861-3873. Peng et al. *Biomacromolecules* 2019; **20**: 3613-3626. Scholz et al. *J Control Release* 2012; **161**: 554-565.]. Among various carriers, non-viral delivery systems have been developed and reported to be more advantageous than the viral delivery system in many aspects [Brito et al. *Adv Genet.* 2015; 89: 179-233]. For example, the large molecular weight branched polyethylenimine (PEI, 25 kDa) is an excellent carrier for plasmid DNA but not for mRNA. However, by decreasing the molecular weight of PEI to 2 kDa, it becomes a more effective carrier of mRNA [Bettinger et al. *Nucleic Acids Res* 2001; **29**: 3882-3891.].

The four-branched histidine-lysine (HK) peptide polymer H2K4b has been shown to be a good carrier of large molecular weight DNA plasmids [Leng et al. *Nucleic Acids Res* 2005; **33**: e40.], but a poor carrier of relatively low molecular weight siRNA [Leng et al. *J Gene Med* 2005; **7**: 977-986.]. Two histidine-rich peptides analogs of H2K4b, namely H3K4b and H3K(+H)4b, were shown to be effective carriers of siRNA [Leng et al. *J Gene Med* 2005; **7**: 977-986. Chou et al. *Biomaterials* 2014; **35**: 846-855.], although H3K(+H)4b appeared to be modestly more effective [Leng et al. *Mol Ther* 2012; **20**: 2282-2290.]. Moreover, the H3K4b carrier of siRNA induced cytokines to a significantly greater degree *in vitro* and *in vivo* than H3K(+H)4b siRNA polyplexes [Leng et al. *Mol Ther* 2012; **20**: 2282-2290.]. Suitable HK polypeptides are described in WO/2001/047496, WO/2003/090719, and WO/2006/060182, the contents of each of which are incorporated herein in their entireties. These polypeptides have a lysine backbone (three lysine residues) where the lysine side chain ε-amino groups and the N-terminus are coupled to various HK sequences. HK polypeptide

carriers can be synthesized by methods that are well-known in the art including, for example, solid-phase synthesis.

It was found that such histidine-lysine peptide polymers ("HK polymers") were surprisingly effective as mRNA carriers, and that they can be used, alone or in combination with liposomes, to provide effective delivery of mRNA into target cells. Similar to PEI and other carriers, initial results suggested HK polymers differ in their ability to carry and release nucleic acids. However, because HK polymers can be reproducibly made on a peptide synthesizer, their amino acid sequence can be easily varied, thereby allowing fine control of the binding and release of mRNAs, as well as the stability of polyplexes containing the HK polymers and mRNA [Chou et al. *Biomaterials* 2014; **35**: 846-855. Midoux et al. *Bioconjug Chem* 1999; **10**: 406-411. Henig et al. *Journal of American Chemical Society* 1999; **121**: 5123-5126.]. When mRNA molecules are admixed with one or more HKP carriers the components self-assemble into nanoparticles.

As described herein, advantageously the HK polymer comprises four short peptide branches linked to a three-lysine amino acid core. The peptide branches consist of histidine and lysine amino acids, in different configurations. The general structure of these histidine-lysine peptide polymers (HK polymers) is shown in Formula I, where R represents the peptide branches and K is the amino acid L-lysine.



Formula I

In Formula I where K is L-lysine and each of R_1, R_2, R_3 and R_4 is independently a histidine-lysine peptide. The R_{1-4} branches may be the same or different in the HK polymers of the invention. When a R branch is "different", the amino acid sequence of that branch differs from each of the other R branches in the polymer. Suitable R branches used in the HK

polymers of the invention shown in Formula I include, but are not limited to, the following R branches R_A – R-$_J$:

R_A = KHKHHKHHKHHKHHKHHKHK– (SEQ ID NO:79)
R_B = KHHHKHHHKHHHKHHHK– (SEQ ID NO:80)
R_C = KHHHKHHHKHHHHKHHHK– (SEQ ID NO:81)
R_D = kHHHkHHHkHHHHkHHHk– (SEQ ID NO:82)
R_E = HKHHHKHHHKHHHHKHHHK– (SEQ ID NO:83)
R_F = HHKHHHKHHHKHHHHKHHHK– (SEQ ID NO:84)
R_G = KHHHHKHHHHKHHHHKHHHHK– (SEQ ID NO:85)
R_H = KHHHKHHHKHHHKHHHHK– (SEQ ID NO:86)
R_I = KHHHKHHHHKHHHKHHHK– (SEQ ID NO:87)
R_J = KHHHKHHHHKHHHKHHHHK– (SEQ ID NO:88)

Specific HK polymers that may be used in the mRNA compositions include, but are not limited to, HK polymers where each of R1, R2, R3 and R4 is the same and selected from R_A – R_J (Table 1).These HK polymers are termed H2K4b, H3K4b, H3K(+H)4b, H3k(+H)4b, H-H3K(+H)4b, HH-H3K(+H)4b, H4K4b, H3K(1+H)4b, H3K(3+H)4b and H3K(1,3+H)4b, respectively. In each of these 10 examples, upper case "K" represents a L-lysine, and lower case "k" represents D-lysine. Extra histidine residues, in comparison to H3K4b, are underlined within the branch sequences. Nomenclature of the HK polymers is as follows:

1) for H3K4b, the dominant repeating sequence in the branches is -HHHK-, thus "H3K" is part of the name; the "4b" refers to the number of branches;

2) there are four -HHHK- motifs in each branch of H3K4b and analogues; the first -HHHK- motif ("1") is closest to the lysine core;

3) H3K(+H)4b is an analogue of H3K4b in which one extra histidine is inserted in the second -HHHK- motif (motif 2) of H3K4b;

4) for H3K(1+H)4b and H3K(3+H)4b peptides, there is an extra histidine in the first (motif 1) and third (motif 3) motifs, respectively; 5) for H3K(1,3+H)4b, there are two extra histidines in both the first and the third motifs of the branches.

Table 1

Polymer	**Branch Sequence**	**Sequence Identifier**
H2K4b	**R_A** = KHKHHKHHKHHKHHKHHKHK-	(SEQ ID NO:79)

	4 3 2 1	
H3K4b	R_B = KHHHKHHHKHHHKHHHK-	(SEQ ID NO:80)
H3K(+H)4b	R_C = KHHHKHHHKHHHKHHHK-	(SEQ ID NO:81)
H3k(+H)4b	R_D = kHHHkHHHkHHHkHHHk-	(SEQ ID NO:82)
H-H3K(+H)4b	R_E = HKHHHKHHHKHHHKHHHK-	(SEQ ID NO:83)
HH-H3K(+H)4b	R_F = HHKHHHKHHHKHHHKHHHK-	(SEQ ID NO:84)
H4K4b	R_G = KHHHHKHHHHKHHHHKHHHK-	(SEQ ID NO:85)
H3K(1+H)4b	R_H = KHHHKHHHKHHHKHHHK-	(SEQ ID NO:86)
H3K(3+H)4b	R_I = KHHHKHHHKHHHKHHHK-	(SEQ ID NO:87)
H3K(1,3+H)4b	R_J = KHHHKHHHKHHHKHHHK-	(SEQ ID NO:88)

Methods well known in the art, including gel retardation assays, heparin displacement assays and flow cytometry can be performed to assess performance of different formulations containing HK polymer plus liposome in successfully delivering mRNA. Suitable methods are described in, for example, Gujrati *et al,* | *Mol. Pharmaceutics* 11:2734−2744 (2014), Pärnaste *et al., Mol Ther Nucleic Acids*. 7: 1–10 (2017).

Detection of mRNA uptake into cells can also be achieved using SmartFlare® technology (Millipore Sigma). These smart flares are beads that have a sequence attached that, when recognizing the RNA sequence in the cell, produce an increase in fluorescence that can be analyzed with a fluorescent microscope.

Other methods include measuring protein expressions from the mRNA - for example, an mRNA encoding luciferase can be used to measure the efficiency of transfection using methods that are well known in the art. See, for example, These were done with luciferase mRNA in the latest publication. (He *et al*, *J Gene Med.* 2021 Feb;23(2):e3295). to demonstrate the efficacy of delivering mRNA using a HKP and liposome formulation.

Synergistic activity of DOTAP and HK carriers in mRNA delivery

The combination of H3K(+H)4b and DOTAP (a cationic lipid) surprisingly was synergistic in its ability to carry mRNA into MDA-MB-231 cells (H3K(+H)4b/liposomes vs liposomes, $P<0.0001$). The combination was about 3-fold and 8-fold more effective as carriers of mRNA than the polymer alone and the cationic lipid carrier, respectively. Not all HK peptides demonstrated the synergistic activity with DOTAP lipid. For example, the

combination of H3K4b and DOTAP was less effective than the DOTAP liposomes as carriers of luciferase mRNA. Besides DOTAP, other cationic lipids that may be used with HK peptides include Lipofectin (ThermoFisher), Lipofectamine (ThermoFisher), and DOSPER (Fig. 22).

The D-isomer of H3k (+H)4b, in which the L-lysines in the branches are replaced with D-lysines, was the most effective polymeric carrier (H3k(+H)4b vs. H3K(+H)4b, $P<0.05$). The D-isomer/ liposome carrier of mRNA was nearly 4-fold and 10-fold more effective than the H3k(+H)4b alone and liposome carrier, respectively. Although the D-H3k(+H)4b/lipid combination was modestly more effective than the L-H3K(+H)4b/lipidmbination, this comparison was not statistically different (Fig. 23).

H3K(+H)4b as a carrier for mRNA among HK polymers

Both H3K4b and H3K(+H)4b may be used as carriers of nucleic acids *in vitro* [Leng *et al. J Gene Med* 2005; 7: 977-986. Chou *et al.*, *Cancer Gene Ther* 2011; 18: 707-716.]. Despite these previous findings, H3K(+H)4b was markedly better as a carrier of mRNA compared to other similar analogues (Table 2).

Table 2

Polymer	Ratio(wt:wt; mRNA:Polymer)	RLU/ug-Protein
H3K(+H)4b	1:4	1532.9±122.9
	1:8	1656.3±202.5
	1:12	1033.4±197
H3k(+H)4b	1:4	1851.6±138.3
	1:8	1787.2±195.2
	1:12	1982.3±210.7
H3K4b	1:4	156.8±41.8
	1:8	62.1±13.2
	1:12	18.1±4.0
H3K(3+H)4b	1:4	61.7±5.7

	1:8	68.7±3.1
	1:12	59.0±7.5
H3K(1+H)4b	1:4	24.3±4.5
	1:8	15.0±3.6
	1:12	7.3±2.5
H-H3K(+H)4b	1:4	1107.5±140.4
	1:8	874.6±65.2
	1:12	676.4±25.7
HH-H3K(+H)4b	1:4	1101.9±106.6
	1:8	832.2±75.3
	1:12	739.8±105.4
H4K4b	1:4	896.4±112.6
	1:8	821.8±115.6
	1:12	522.4±69.2
H3K(1,3+H)4b	1:4	518.3±134.7
	1:8	427.7±18.1
	1:12	378±5.2
H2K4b	1:4	546.7±70.1
	1:8	132.3±58.5
	1:12	194.7±18.4

Especially, it has higher mRNA transfection efficiency than H3K4b in various weight:weight (HK:mRNA) ratios. At a 4:1 ratio, luciferase expression was 10-fold higher with H3K(+H)4b than H3K4b in MDA-MB-231 cells without significant cytotoxicity. Moreover, the buffering capacity does not seem to be an essential factor in their transfection differences since the percent of histidines (by weight) in H3K4b and H3K(+H)4b is 68.9 and 70.6 %, respectively.

Gel retardation assays show that the electrophoretic mobility of mRNA was delayed by the HK polymers. The retardation effect increased with higher peptide to mRNA weight

ratios. However, mRNA was completely retarded in 2:1 ratio of H3K(+H)4b, whereas it was not completely retarded by H3K4b. This suggested that H3K(+H)4b could form a more stable polyplex, which was advantageous for its ability to be a suitable carrier for mRNA delivery.

Further confirmation that the H3K(+H)4b peptide binds more tightly to mRNA was demonstrated with a heparin-displacement assay. Various concentrations of heparin was added into the polyplexes formed with mRNA and HK and it was observed that, particularly at the lower concentrations of heparin, mRNA was released by the H3K4b polymer more readily than the H3K(+H)4b polymer. These data suggest H3K(+H)4b could bind to mRNA and form a more stable polyplex than H3K4b.

With the mRNA labeled with cyanine-5, the uptake of H3K4b and H3K(+H)4b polyplexes into MDA-MB-231 cells was compared using flow cytometry. At different time points (1, 2, and 4 h), the H3K(+H)4b polyplexes were imported into the cells more efficiently than H3K4b polyplexes (data not shown). Similar to these results, fluorescent microscopy indicated that H3K(+H)4b polyplexes localized within the acidic endosomal vesicles significantly more than H3K4b polyplexes (H3K4b vs. H3K(+H)4b, $P<0.001$). Interestingly, irregularly-shaped H3K4b polyplexes, which did not overlap endocytic vesicles, were likely extracellular and were not observed with H3K(+H)4b polyplexes.

It is known both that HK polymers and cationic lipids (i.e., DOTAP) significantly and independently increase transfection with plasmids [Chen et al. *Gene Ther* 2000; **7**: 1698-1705.]. Consequently, whether these lipids together with HK polymers enhanced mRNA transfection was investigated. Notably, the H3K(+H)4b and H3k(+H)4b carriers were significantly better carriers of mRNA than the DOTAP liposomes. The combination of H3K(+H)4b and DOTAP lipid was synergistic in the ability to carry mRNA into MDA-MB-231 cells. The combination was about 3-fold and 8-fold more effective as carriers of mRNA than the polymer alone and the liposome carrier, respectively (H3K(+H)4b/lipid vs. liposomes or H3K(+H)4b). Notably, not all HK peptides demonstrated improved activity with DOTAP lipid. The combination of H3K4b and DOTAP carriers was less effective than the DOTAP liposomes as carriers of luciferase mRNA. The combination of DOTAP and

H3K(+H)4b carriers were found to be synergistic in their ability to carry mRNA into cells [He et al. *J Gene Med.* 2020 Nov 10:e3295].

Administration of mRNA Vaccines

The vaccine formulations described herein may be administered to subjects, including human subjects, by any mode of administration that is conventionally used to administer vaccines. Thus the compositions can be in the form of an aerosol, dispersion, solution, or suspension and can be formulated for inhalation, intramuscular, oral, sublingual, buccal, parenteral, nasal, subcutaneous, intradermal, or topical administration. The term parenteral as used herein includes percutaneous, subcutaneous, intravascular (e.g., intravenous), intramuscular, or intrathecal injection or infusion techniques and the like.

As used herein, an effective dose of a vaccine is the dose required to produce a protective immune response in the subject to whom the vaccine is administered. A protective immune response in the present context is one that prevents or ameliorates disease in a subject challenged with 2019-nCoV. Methods of running clinical trials in multiple subjects to determine whether a vaccine produces a protective immune response are well known in the art.

The vaccine may be administered one or more times. An initial measurement of an immune response to the vaccine may be made by measuring production of antibodies in the subject receiving the vaccine. Methods of measuring antibody production in this manner are also well known in the art. is that dose required to prevent, inhibit the occurrence, or treat (alleviate a symptom to some extent, preferably all of the symptoms) of a disease state. The pharmaceutically effective dose depends on the type of disease, the composition used, the route of administration, the type of mammal being treated, the physical characteristics of the specific mammal under consideration, concurrent medication, and other factors that those skilled in the medical arts will recognize Generally, an amount between 0.1 mg/kg and 100 mg/kg body weight/day of active ingredients is administered dependent upon potency of the formulated composition.

However, their application has until recently been restricted by the instability and inefficient *in vivo* delivery of mRNA. The methods described herein provide methods of making and using mRNA vaccines with a polypeptide delivery system and a polypeptide nanoparticle delivery system.

The methods described herein may be used in clinical applications of the mRNA include prophylactic and therapeutic vaccines against various diseases, especially infectious diseases.

Example 1: mRNA synthesis:

CleanCap SARS-COV 2 spike protein (S1P and S2P) and CleanCap Enhanced Green Fluorescent Protein(EGFP) expressed mRNA were synthesized and CleanCap SARS-COV 2 spike protein (S6P) was manufactured by TriLink BioTechnologies, Inc, (San Diego, CA) and purified by selective binding of dsRNA to cellulose in an ethanol-containing bufferA chromatography buffer containing 10mM HEPES (pH 7.2), 0.1mM EDTA, 125mM NaCl, 16% ethanol and cellulose fibers were added to microcentrifuge spin columns along with mRNA products and centrifuged. Almost 90% of dsRNA could be removed after this procedure (Baiersdorfer et.al, 2019). Contaminants could be also eliminated using FPLC and HPLC (Kariko et.al, 2011).

Example 2: Peptides (HK polymers) preparation

The HK peptide polymers were synthesized on a Rainin Voyager synthesizer (Tucson, AZ) by the biopolymer core facility at the University of Maryland.

Example 3: In-vitro transfection of mRNA

To verify the proper protein expression of mRNA, EGFP mRNA and SARS-COV 2 SP, mRNAs were transiently transfected into human embryonic kidney 293T cells (293T cells). Briefly, 4.8 X 10^5 cells were plated into a 6 well plate containing 2ml of DMEM(10% fetal bovine serum and 1% Penicillin-Streptomycin (ThermoFisher Scientific)). After 24hr, when the cells were 70-90% confluent, mRNAs were transfected into 293T cells using

Lipofectamine MessengerMAX Transfection Reagent (ThermoFisher Scientific) according to the provided protocol. 293T celsl were cultured for two day for *in vitro* protein expression measurement.

Two HK based polymers associated with cationic liposome DOTAP (1,2-dioleoyl-3-trimethylammonium-propane, Roche), MC3 (DLin-MC3-DMA, ChemScene), and PLA(Polylactic acid) were examined for their ability to carry an EGFP mRNA and SARS-COV 2 SP mRNA into human embryonic kidney 293 cells(HEK 293 cells). Briefly, 4.8 X 10^5 cells were plated into a 6 well plate containing 2ml of DMEM(10% fetal bovine serum and Penicillin-Streptomycin (ThermoFisher Scientific)). After 24hr, when the cell were 70-90% confluent, two HK based polymer formulations with EGFP and SPs mRNA added into each well. 293T cell were cultured for two day for In-vitro protein expression measurement.

Example 4: In-vitro protein expression measurement

Immunofluorescence analysis: After two days of transfection, protein expression was measured by immunofluorescence imaging using a Cytation5 Cell Imaging Multi-Mode Reader (Biotek, Winooski, VT) .

Cell lysate preparation: After two days of transfection, culture media was aspirated and cells were washed on ice with ice-cold PBS. Ice-cold lysis buffer(RIPA, ThermoFisher Scientific) with protease inhibitor(ThermoFisher Scientific) was added and cells were incubated for 30 minutes at 4°C. Cells were harvested using a cell scraper and lysed by sonication. Centrifugation at 10,000g for 20 minutes at 4°C pelleted cell debris, and the supernatant was transferred to a fresh microcentrifuge tube. The protein concentration of the lysate was determined by Bradford or BCA protein assay for Western blot.

Western blot: Briefly, in each well of a gel, 20-50ug of protein was with 4X SDS sample buffer(ThermoFisher Scientific), 10X Reducing buffer(ThermoFisher Scientific), and additional ddH2O(ThermoFisher Scientific) with a total loading volume of 25µl/well. The mixture was denatured by heating at 95°C for 5 minutes and cooled to room temperature and centrifuged before loading onto a NuPAGE™ 4 to 12%, Bis-Tris gel(ThermoFisher Scientific). After electrophoretic separation, the gel was removed from the cassette and

transferred using an iBlot™ 2 Dry Blotting System (ThermoFisher Scientific). The transfer membrane was blocked with 5% fat-free milk powder in TBST for 1hr at RT, incubated with primary antibody for overnight at 4°C, washed three times with TBST(0.05% Tween20 in TBS) buffer, and incubated with secondary antibody, which is HRP conjugated Mouse IgG (H+L) Secondary Antibody (ThermoFisher Scientific, A24512) for 1hr at RT. Transfer membrane was developed by Pierce ECL Western Blotting Substrate (ThermoFisher Scientific) and was imaged using chemiluminescent imaging system.

Example 5: In vivo formulation preparation

1. PNI-Genvoy LNP formulation: Lipid nanoparticles were formulated using the GenVoy Platform with PNI NanoAssemblr (Precision NanoSystems, Vancouver, British Columbia, Canada) as the positive control in both *in vitro* and *in vivo* assays.

2. HKP(+H) formulation: HKP(+H) stock solution (10 mg/mL) was prepared in nuclease free water. A concentrated stock solution was diluted to 4 mg/mL in water. mRNA working solution (1 mg/mL) was prepared in 1 mM citrate buffer (pH 6.0). The mRNA/HKP(+H) polyplex was formed by mixing equal volumes of 4 mg/mL HKP(+H) and 1 mg/mL mRNA. The mass ratio of HKP(+H) to mRNA is 4:1. The mRNA/HKP(+H) polyplex was incubated for 30 min at room temperature before use. The size of each peptide-based polyplexwas determined with the Zetasizer(Malvern Panalytical) prior to transfection or injection.

3. HKP(+H)/DOTAP formulation (post-mixed DOTAP): HKP(+H) stock solution (10 mg/mL) was prepared in nuclease free water. A concentrated stock solution was diluted to 4 mg/mL in water. DOTAP (Sigma-Aldrich) is 1 mg/mL in aqueous buffered solution. mRNA working solution (1 mg/mL) was prepared in 1 mM citrate buffer (pH 6.0). First the mRNA/HKP(+H) polyplex was formed by mixing equal volumes of 4 mg/mL HKP(+H) and 1 mg/mL mRNA. The mRNA/HKP(+H) polyplex was incubated for 30 min at room temperature. Next, the same volume of DOTAP to HKP(+H) solution was added to the mRNA/HKP(+H) polyplex. The mass ratio of HKP(+H)/DOTAP to mRNA was 4:1:1. The

mRNA/HKP(+H)/DOTAP nanoparticle was incubated for 30 min at room temperature before use.

4. HKP(+H)/MC3 or HKP(+H)/DOTAP formulations (pre-mixed MC3 or DOTAP): HKP(+H) stock solution (10 mg/mL) was prepared in nuclease-free water. A concentrated stock solution was diluted to 4 mg/mL in water. DOTAP or MC3 is 1 mg/mL in aqueous buffered solution. mRNA working solution (1 mg/mL) was prepared in 1 mM citrate buffer (pH 6.0). Equal volumes of HKP(+H) and MC3 were pre-mixed at a 4:1 mass ratio and the same volume of mRNA to HKP(+H) solution was added to a pre-mixed HKP(+H)/MC3. The mRNA/HKP(+H)/MC3 nanoparticle was formed by mixing pre-mixed 4 mg/mL HKP(+H)/ 1mg/mL MC3 and 1 mg/mL mRNA. The mass ratio of HKP(+H)/MC3 to mRNA is 4:1:1. The mRNA/HKP(+H)/MC3 nanoparticle was incubated for 30 min at room temperature before use.

5. HKP(+H)/PLA NP formulation: HKP(+H) stock solution (10 mg/mL) was prepared in nuclease free water. A concentrated stock solution was diluted to 4 mg/mL in water. A poly-L-Lactic Acid (PLA) nanoparticle (5mg/mL) was prepared in water. mRNA working solution (1 mg/mL) was prepared in 1 mM citrate buffer (pH 6.0). Equal volumes of HKP(+H) and mRNA were mixed at a 4:1 mass ratio. The mRNA/HKP(+H) polyplex was incubated for 30 min at room temperature, and then the same volume of PLA nanoparticle to HKP(+H) solution was added to the mRNA/HKP(+H) nanoparticle so that the mRNA/HKP(+H) polyplex was adsorbed on the surface of the PLA nanoparticle. The mass ratio of HKP(+H)/PLA to mRNA was 4:5:1. The mRNA/HKP(+H)/PLA nanoparticle was incubated for 30 min at room temperature before use.

Example 6: In vivo animal model and injection

An *in vivo* study was performed at Antibody and Immunoassay Consultants (AIC) Inc (Gaithersburg, MD). Briefly,8 6-8 week old female BALB/C were randomized into each group, 4 mice per each group and injected intramuscularly into the right flank with 30μg of EGFP with different formulations. With the same formulations mentioned above, CleanCap

SARS-COV 2 spike protein (S6P) polyplexes were prepared for *in vivo* analysis and antibody titer measurement and binding.

On day 28, a second injection for boosting was made and on day 35 serum was collected and analyzed by immunoassay (ELISA) for measurement of antibody titer.

Results

Immunofluorescence Analysis

One week after fabrication, formulations delivered to HEK 293T cells (Figure 2). All HKP(H) groups proved the possibility to carry mRNA into cells in vitro. Among HKP(H) groups, HKP(H) alone group showed higher EGFP expression than others and PNI Genvoy (positive control).

Western Blot Analysis

Next Western Blot assay performed and results (Figure 3) showed the expressed spike protein on the cell surface from S1P D614G mRNA delivered group was underwent shedding after localization to cell membrane. Only S2 domain of spike protein remained on the cell surface. The band size in the Western Blot(70-80 KDa) supported this hypothesis. However the spike protein from S2P D614G showed intact size of the spike protein(180 KDa).

WHAT IS CLAIMED IS:

1. A composition comprising an effective amount of a messenger ribonucleic acid (mRNA) that comprises an open reading frame (ORF) encoding a 2019-nCoV protein or protein fragment formulated in a pharmaceutically acceptable carrier, wherein said mRNA and said carrier form a polymeric nanoparticle.

2. The composition of claim 1, wherein the 2019-nCoV protein is a spike (S) protein or spike (S) subunit.

3. The composition of claim 2, wherein the ORF is selected from the group consisting of mRNA molecules encoding a peptide selected from SEQ ID 14 to 31.

4. The composition according to claim 2, wherein said ORF is selected from the group consisting of SEQ ID No.32 to 49.

5. The composition of claim 2, wherein the ORF comprises an mRNA encoding a peptide selected from SEQ ID No.50 to 55.

6. The composition of claim 4, wherein said ORF is selected from the group consisting of SEQ ID No.91 to 96.

7. The composition of claim 2, wherein the ORF comprises an mRNA encoding a peptide selected from SEQ ID No.56 to 66.

8. The composition of claim 6, wherein said ORF is selected from the group consisting of SEQ ID No.97 to 107.

9. The composition of claim 2 wherein the ORF is selected from the group consisting of mRNA molecules encoding the amino acid sequence selected from SEQ ID No. 11, 13, 68, 70, 72, 74 and 90

10. The composition of claim 8 wherein said ORF is selected from the group consisting of SEQ ID No.10, 12, 67, 69, 71, 73 and 89.

11. The composition of claim 2 wherein said ORF is selected from the group consisting of SEQ ID No. 1, 2, 3, and 4.

12. The composition of any preceding claim, further comprising an mRNA comprising all or part of orf1a and/or orf1b, wherein said mRNA encodes a 2019-nCoV non-structural protein polymerase and/or RdRp.

13. The composition of claim 12 wherein said ORF is an mRNA selected from the group consisting of SEQ ID NOs. 5 and 6.

14. The composition of claim 1 wherein said ORF encodes all or part of the 2019 nCoV encoding envelope (E), membrane (M), and/or nucleocapsid (N) protein.

15. The composition of any preceding claim, further comprising one or more mRNA sequences encoding all or part of the 2019 nCoV encoding envelope (E), membrane (M), and/or nucleocapsid (N) protein.

16. The composition of claim 14 or 15, wherein said ORF is selected from the group consisting of SEQ ID No.7, 8, and 9.

17. The composition of any preceding claim, wherein the mRNA is a chemically modified mRNA sequence comprising one or more N1-methyl-pseudouridine and/or pseudouridine residues.

18. The composition according to claim 17, wherein all uridine residues in the mRNA are N1-methyl pseudouridine or pseudouridine.

19. The composition of any preceding claim, wherein the mRNA molecule further comprises a 3'-UTR, a 5'-UTR and optionally further comprises additional elements that (a) stabilize the composition and (b) enhance expression of the polypeptide encoded by said ORF.

20. The composition of claim 19, wherein the 5'-UTR comprises an m7G cap structure and a start codon.

21. The composition of claim 19 or 20, wherein the 3'-UTR comprises a stop codon and a polyA tail.

22. A method of making a composition according to any one of claims 1-21, comprising expressing said ORF in a linear *in vitro* transcription (IVT) system.

23. A method of making a compositions of any one of claims 1-21, comprising expressing said ORF expression using a plasmid DNA (pDNA) vector delivery system.

24. The composition of any one of claims 1-21, wherein the composition comprises more than one ORF and encodes at least two polypeptides.

25. The composition of any one of claims 1-21 or 24, wherein the composition further comprises a polymeric nanoparticle carrier comprising a Histidine-Lysine co-polymer (HKP).

26. The composition of claim 25, wherein the HKP is selected from the group consisting of polymers having side chains selected from SEQ ID NO:75-88.

27. A method of making a composition according to claim 25 or claims 26, comprising contacting a composition comprising said ORF with an HKP, said mRNA and said HKP self-assemble into nanoparticles.

28. The composition of any of claims 1-21 or 24-26, wherein said nanoparticle further comprises a lipid, wherein said lipid optionally is a cationic lipid.

29. A method of treating a subject with a 2019-nCoV infection comprising administering to said subject a pharmaceutically effective amount of a composition according to any one of claims 1-21, 24-26 or28.

30. The method of claim 29, wherein the composition is atomized by a nebulizer inhalation system prior to or during administration.

31. The method of claim 30, wherein the nebulizer system is a portable nebulizer for whole respiratory tract drug delivery.

32. The method of claim 29, wherein the composition is administered by subcutaneous injection.

33. The method of claim 29, wherein the composition is administered by intramuscular injection.

34. The method of claim 29, wherein the composition is administered by intraperitoneal injection (i.p).

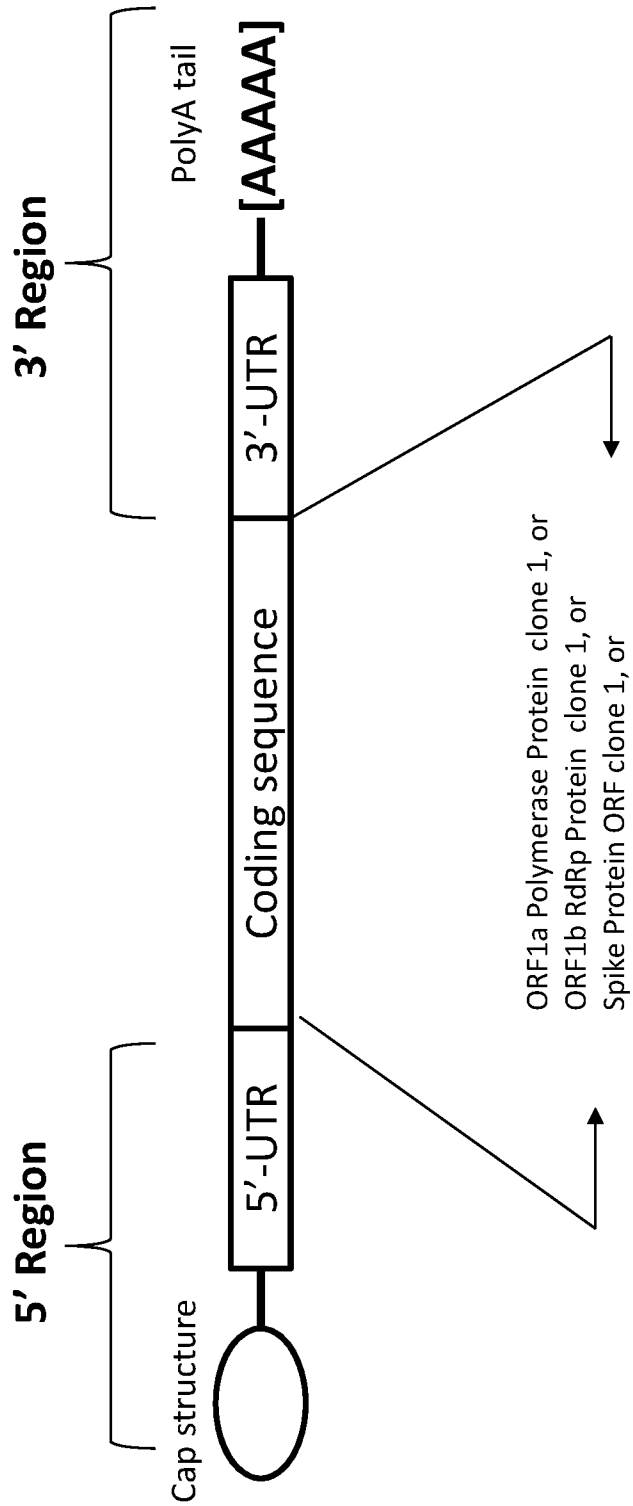


Figure 1. Schematic representation of optimized mRNA vaccine expression structure

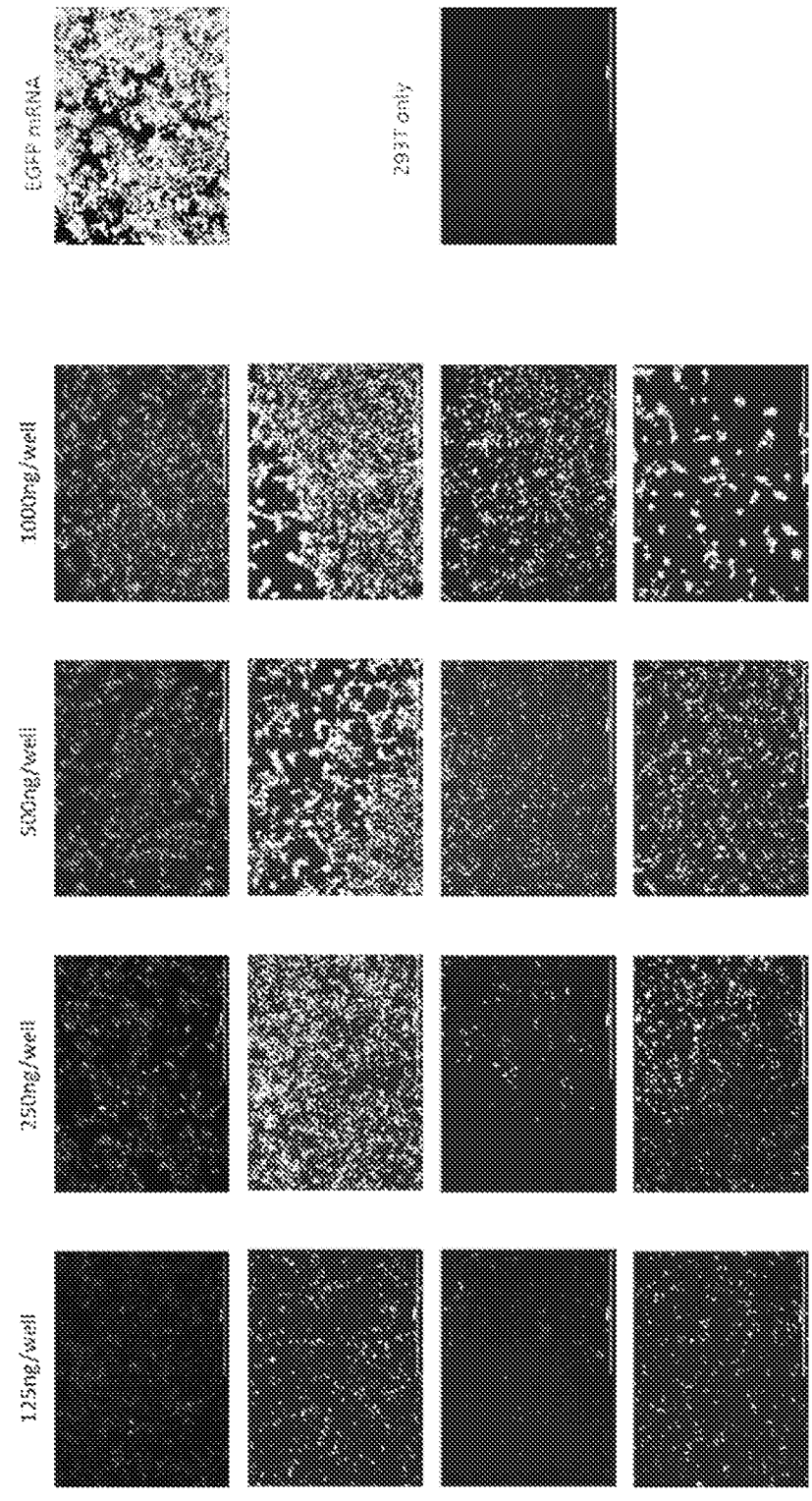


Figure 2. Immunofluorescence analysis result of Spike Protein expression mRNA transient transfection. Transfection condition : ThermoFisher Scientific Lipofectamine MessengerMAX Transfection Reagent protocol(Pin. No. MAN0010803 Rev D.0)

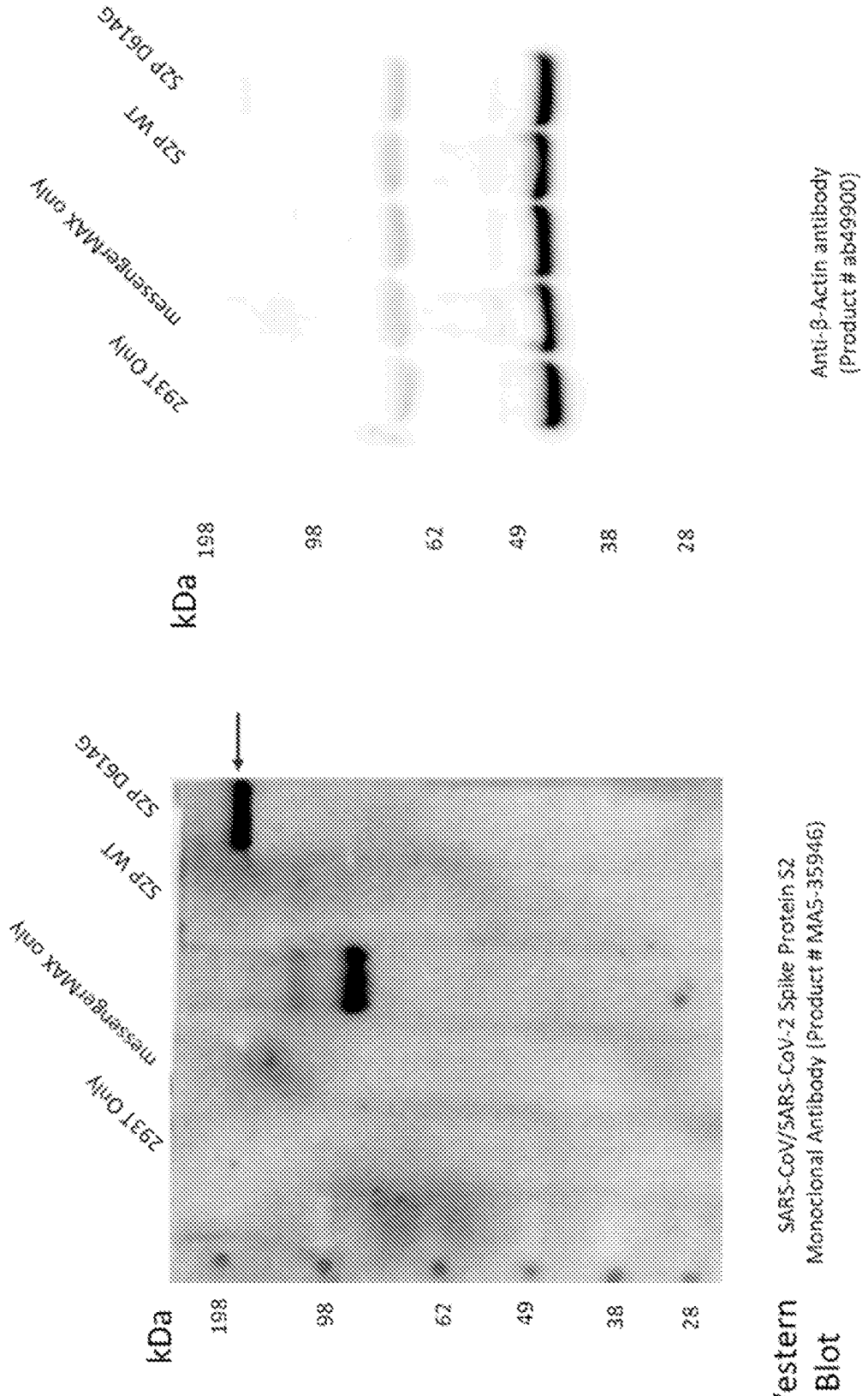


Figure 3. Western Blot result of Spike Protein expression mRNA transient transfection. Trasfection condition : ThermoFisher Scientific Lipofectamine MessengerMAX Transfection Reagent protocol(Pin. No. MAN0010803 Rev D.0). Primary antibody : SARS-CoV/SARS-CoV-2 Spike Protein S2 Monoclonal Antibody (ThermoFisher Scientific, MA5-35946). Secondary antibody : Mouse IgG (H+L) Secondary Antibody (ThermoFisher Scientific, A24512). Loading control : HRP Anti-beta Actin antibody (abcam, ab49900)

Figure 4:

Spike Protein ORF clone 1 (MN996531: viral cDNA: 21523-23040) SEQ ID NO:1

atgttctt gttaacaact 21541
aaacgaacaa tgtttgtttt tcttgtttta ttgccactag tctctagtca gtgtgttaat cttacaacca gaactcaatt
acccccctgca tacactaatt ctttcacacg tggtgtttat taccctgaca aagttttcag atcctcagtt ttacattcaa
ctcaggactt gttcttacct ttcttttcca atgttacttg gttccatgct atacatgtct ctgggaccaa tggtactaag
aggtttgata accctgtcct accatttaat gatggtgttt attttgcttc cactgagaag tctaacataa taagaggctg
gatttttggt actactttag attcgaagac ccagtcccta cttattgtta ataacgctac taatgttgtt attaaagtct
gtgaatttca attttgtaat gatccatttt tgggtgttta ttaccacaaa aacaacaaaa gttggatgga aagtgagttc
agagtttatt ctagtgcgaa taattgcact tttgaatatg tctctcagcc ttttcttatg gaccttgaag gaaaacaggg
taatttcaaa aatcttaggg aatttgtgtt taagaatatt gatggttatt ttaaaatata ttctaagcac acgcctatta
atttagtgcg tgatctccct cagggttttt cggctttaga accattggta gatttgccaa taggtattaa catcactagg
tttcaaactt tacttgcttt acatagaagt tatttgactc ctggtgattc ttcttcaggt tggacagctg gtgctgcagc
ttattatgtg ggttatcttc aacctaggac ttttctatta aaatataatg aaaatggaac cattacagat gctgtagact
gtgcacttga ccctctctca gaaacaaagt gtacgttgaa atccttcact gtagaaaaag gaatctatca aacttctaac
tttagagtcc aaccaacaga atctattgtt agatttccta atattacaaa cttgtgccct tttggtgaag tttttaacgc
caccagattt gcatctgttt atgcttggaa caggaagaga atcagcaact gtgttgctga ttattctgtc ctatataatt
ccgcatcatt ttccactttt aagtgttatg gagtgtctcc tactaaatta aatgatctct gctttactaa tgtctatgca
gattcatttg taattagagg tgatgaagtc agacaaatcg ctccagggca aactggaaag attgctgatt ataattataa
attaccagat gattttacag gctgcgttat agcttggaat tctaacaatc ttgattctaa ggttggtggt aattataatt
acctgtatag attgtttagg aagtctaatc tcaaaccttt tgagagagat atttcaactg aaatctatca ggccggtagc
acaccttgta atggtgttga aggttttaat tgttactttc ctttacaatc atatggtttc

Spike Protein ORF clone 2, (MN996531: viral cDNA: 23041-23740) SEQ ID NO:2

23041
caacccacta atggtgttgg ttaccaacca tacagagtag tagtactttc ttttgaactt ctacatgcac cagcaactgt
ttgtggacct aaaaagtcta ctaatttggt taaaaacaaa tgtgtcaatt tcaacttcaa tggtttaaca ggcacaggtg
ttcttactga gtctaacaaa aagtttctgc ctttccaaca atttggcaga gacattgctg acactactga tgctgtccgt
gatccacaga cacttgagat tcttgacatt acaccatgtt cttttggtgg tgtcagtgtt ataacaccag gaacaaatac
ttctaaccag gttgctgttc tttatcagga tgttaactgc acagaagtcc ctgttgctat tcatgcagat caacttactc
ctacttggcg tgtttattct acaggttcta atgtttttca aacacgtgca ggctgtttaa taggggctga acatgtcaac
aactcatatg agtgtgacat acccattggt gcaggtatat gcgctagtta tcagactcag actaattctc
ctcggcgggc acgtagtgta gctagtcaat ccatcattgc ctacactatg tcacttggtg cagaaaattc agttgcttac
tctaataact ctattgccat acccacaaat tttactatta gtgttaccac agaaattcta ccagtgtcta

Spike Protein ORF clone 3 (MN996531: viral cDNA: 23740-24540) SEQ ID NO:3

23740
tgaccaagac atcagtagat tgtacaatgt acatttgtgg tgattcaact gaatgcagca atcttttgtt gcaatatggc
agtttttgta cacaattaaa ccgtgcttta actggaatag ctgttgaaca agacaaaaac acccaagaag tttttgcaca
agtcaaacaa atttacaaaa caccaccaat taaagatttt ggtggtttta atttttcaca aatattacca gatccatcaa
aaccaagcaa gaggtcattt attgaagatc tacttttcaa caaagtgaca cttgcagatg ctggcttcat caaacaatat
ggtgattgcc ttggtgatat tgctgctaga gacctcattt gtgcacaaaa gtttaacggc cttactgttt tgccaccttt
gctcacagat gaaatgattg ctcaatacac ttctgcactg ttagcgggta caatcacttc tggttggacc tttggtgcag
gtgctgcatt acaaatacca tttgctatgc aaatggctta taggtttaat ggtattggag ttacacagaa tgttctctat
gagaaccaaa aattgattgc caaccaattt aatagtgcta ttggcaaaat tcaagactca ctttcttcca cagcaagtgc
acttggaaaa cttcaagatg tggtcaacca aaatgcacaa gctttaaaca cgcttgttaa acaacttagc tccaattttg
gtgcaatttc aagtgtttta aatgatatcc tttcacgtct tgacaaagtt gaggctgaag tgcaaattga taggttgatc

Spike Protein ORF clone 4 (MN996531: viral cDNA: 24540-25371) SEQ ID NO:4

24540
acaggcagac ttcaaagttt gcagacatat gtgactcaac aattaattag agctgcagaa atcagagctt
ctgctaatct tgctgctact aaaatgtcag agtgtgtact tggacaatca aaaagagttg atttttgtgg aaagggctat
catcttatgt ccttccctca gtcagcacct catggtgtag tcttcttgca tgtgacttat gtccctgcac aagaaaagaa
cttcacaact gctcctgcca tttgtcatga tggaaaagca cactttcctc gtgaaggtgt ctttgtttca aatggcacac
actggtttgt aacacaaagg aatttttatg aaccacaaat cattactaca gacaacacat ttgtgtctgg taactgtgat
gttgtaatag gaattgtcaa caacacagtt tatgatcctt tgcaacctga attagactca ttcaaggagg agttagataa
atattttaag aatcatacat caccagatgt tgatttaggt gacatctctg gcattaatgc ttcagttgta aacattcaaa
aagaaattga ccgcctcaat gaggttgcca agaatttaaa tgaatctctc atcgatctcc aagaacttgg
aaagtatgag cagtatataa aatggccatg gtacatttgg ctaggtttta tagctggctt gattgccata gtaatggtga
caattatgct ttgctgtatg accagttgct gtagttgtct caagggctgt tgttcttgtg gatcctgctg caaatttgat
gaagacgact ctgagccagt gctcaaagga gtcaaattac attacacata a

ORF1a Polymerase Protein clone 1 (MN996531: viral cDNA: 6253-7140) SEQ ID NO:5

6253
gccacgta taaaccaaat acctggtgta tacgttgtct ttggagcaca aaaccagttg aaacatcaaa ttcgtttgat
gtactgaagt cagaggacgc gcagggaatg gataatcttg cctgcgaaga tctaaaacca gtctctgaag
aagtagtgga aaatcctacc atacagaaag acgttcttga gtgtaatgtg aaaactaccg aagttgtagg
agacattata cttaaaccag caaataatag tttaaaaatt acagaagagg ttggccacac agatctaatg gctgcttatg
tagacaattc tagtcttact attaagaaac ctaatgaatt atctagagta ttaggtttga aaacccttgc tactcatggt
ttagctgctg ttaatagtgt cccttgggat actatagcta attatgctaa gccttttctt aacaaagttg ttagtacaac
tactaacata gttacacggt gtttaaaccg tgtttgtact aattatatgc cttatttctt tactttattg ctacaattgt
gtacttttac tagaagtaca aattctagaa ttaaagcatc tatgccgact actatagcaa agaatactgt taagagtgtc
ggtaaatttt gtctagaggc ttcatttaat tatttgaagt cacctaattt ttctaaactg ataaatatta taatttggtt

tttactatta agtgtttgcc taggttcttt aatctactca accgctgctt taggtgtttt aatgtctaat ttaggcatgc cttcttactg tactggttac agagaaggct atttgaactc tactaatgtc actattgcaa cctactgtac tggttctata ccttgtagtg tttgtcttag tggtttagat

ORF1b RdRp Protein clone 1 (MN996531: viral cDNA: 15253-16153) SEQ ID NO:6

15253
tagaaaac cctcacctta tgggttggga ttatcctaaa tgtgatagag ccatgcctaa catgcttaga attatggcct cacttgttct tgctcgcaaa catacaacgt gttgtagctt gtcacaccgt ttctatagat tagctaatga gtgtgctcaa gtattgagtg aaatggtcat gtgtggcggt tcactatatg ttaaaccagg tggaacctca tcaggagatg ccacaactgc ttatgctaat agtgttttta acatttgtca agctgtcacg gccaatgtta atgcactttt atctactgat ggtaacaaaa ttgccgataa gtatgtccgc aatttacaac acagacttta tgagtgtctc tatagaaata gagatgttga cacagacttt gtgaatgagt tttacgcata tttgcgtaaa catttctcaa tgatgatact ctctgacgat gctgttgtgt gtttcaatag cacttatgca tctcaaggtc tagtggctag cataaagaac tttaagtcag ttctttatta tcaaaacaat gtttttatgt ctgaagcaaa atgttggact gagactgacc ttactaaagg acctcatgaa ttttgctctc aacatacaat gctagttaaa cagggtgatg attatgtgta ccttccttac ccagatccat caagaatcct aggggccggc tgttttgtag atgatatcgt aaaaacagat ggtacactta tgattgaacg gttcgtgtct ttagctatag atgcttaccc acttactaaa catcctaatc aggagtatgc tgatgtcttt catttgtact tacaatacat aagaaagcta catgatgagt taacaggaca catgttagac atgtattctg ttatgcttac taa

M Protein ORF (MN996531: viral cDNA: 26510-27178) SEQ ID NO:7

a tggcagattc 26521 caacggtact attaccgttg aagagcttaa aaagctcctt gaacaatgga acctagtaat aggtttccta ttccttacat ggatttgtct tctacaattt gcctatgcca acaggaatag gtttttgtat ataattaagt taattttcct ctggctgtta tggccagtaa ctttagcttg ttttgtgctt gctgctgttt acagaataaa ttggatcacc ggtggaattg ctatcgcaat ggcttgtctt gtaggcttga tgtggctcag ctacttcatt gcttctttca gactgtttgc gcgtacgcgt tccatgtggt cattcaatcc agaaactaac attcttctca acgtgccact ccatggcact attctgacca gaccgcttct agaaagtgaa ctcgtaatcg gagctgtgat ccttcgtgga catcttcgta ttgctggaca ccatctagga cgctgtgaca tcaaggacct gcctaaagaa atcactgttg ctacatcacg aacgctttct tattacaaat tgggagcttc gcagcgtgta gcaggtgact caggttttgc tgcatacagt cgctacagga ttggcaacta taaattaaac acagaccatt ccagtagcag tgacaatatt gctttgcttg tacagtaa

N Protein ORF (MN996531: viral cDNA: 28261-29520) SEQ ID NO:8

28261
atgtctgata atggacccca aaatcagcga aatgcacccc gcattacgtt tggtggaccc tcagattcaa ctggcagtaa ccagaatgga gaacgcagtg gggcgcgatc aaaacaacgt cggccccaag gtttacccaa taatactgcg tcttggttca ccgctctcac tcaacatggc aaggaagacc ttaaattccc tcgaggacaa ggcgttccaa ttaacaccaa tagcagtcca gatgaccaaa ttggctacta ccgaagagct accagacgaa ttcgtggtgg tgacggtaaa atgaaagatc tcagtccaag atggtatttc tactacctag gaactgggcc agaagctgga cttccctatg gtgctaacaa agacggcatc atatgggttg caactgaggg agccttgaat

acaccaaaag atcacattgg cacccgcaat cctgctaaca atgctgcaat cgtgctacaa cttcctcaag
gaacaacatt gccaaaaggc ttctacgcag aagggagcag aggcggcagt caagcctctt ctcgttcctc
atcacgtagt cgcaacagtt caagaaattc aactccaggc agcagtaggg gaacttctcc tgctagaatg
gctggcaatg gcggtgatgc tgctcttgct ttgctgctgc ttgacagatt gaaccagctt gagagcaaaa
tgtctggtaa aggccaacaa caacaaggcc aaactgtcac taagaaatct gctgctgagg cttctaagaa
gcctcggcaa aaacgtactg ccactaaagc atacaatgta acacaagctt tcggcagacg tggtccagaa
caaacccaag gaaattttgg ggaccaggaa ctaatcagac aaggaactga ttacaaacat tggccgcaaa
ttgcacaatt tgcccccagc gcttcagcgt tcttcggaat gtcgcgcatt ggcatggaag tcacaccttc
gggaacgtgg ttgacctaca caggtgccat caaattggat gacaaagatc caaatttcaa agatcaagtc
attttgctga ataagcatat tgacgcatac aaaacattcc caccaacaga gcctaaaaag gacaaaaaga
agaaggctga tgaaactcaa gccttaccgc agagacagaa gaaacagcaa actgtgactc ttcttcctgc
tgcagatttg gatgatttct ccaaacaatt gcaacaatcc atgagcagtg ctgactcaac tcaggcctaa

E Protein ORF (MN996531: viral cDNA: 26232-26459) SEQ ID NO:9

26232
atgtactca ttcgtttcgg aagagacagg tacgttaata gttaatagcg tacttctttt tcttgctttc gtggtattct
tgctagttac actagccatc cttactgcgc ttcgattgtg tgcgtactgc tgcaatattg ttaacgtgag tcttgtaaaa
ccttcttttt acgtttactc tcgtgttaaa aatctgaatt cttctagagt tcctgatctt ctggtctaa

Figure 5

D614G Mutated Spike Protein ORF SEQ ID NO:10

atgtttgtttttcttgttttattgccactagtctctagtcagtgtgttaatcttacaaccagaactcaattacccccctgcatacactaat
tctttcacacgtggtgtttattaccctgacaaagttttcagatcctcagttttacattcaactcaggacttgttcttacctttcttttcc
aatgttacttggttccatgctatacatgtctctgggaccaatggtactaagaggtttgataaccctgtcctaccatttaatgatgg
tgtttattttgcttccactgagaagtctaacataataagaggctggatttttggtactactttagattcgaagacccagtccctact
tattgttaataacgctactaatgttgttattaaagtctgtgaatttcaattttgtaatgatccatttttgggtgtttattaccacaaaaa
caacaaaagttggatggaaagtgagttcagagtttattctagtgcgaataattgcacttttgaatatgtctctcagccttttcttat
ggaccttgaaggaaaacagggtaatttcaaaaatcttagggaatttgtgtttaagaatattgatggttatttaaaatatattctaa
gcacacgcctattaatttagtgcgtgatctccctcagggtttttcggctttagaaccattggtagatttgccaataggtattaaca
tcactaggtttcaaacttacttgctttacatagaagttatttgactcctggtgattcttcttcaggttggacagctggtgctgcag
cttattatgtgggttatcttcaacctaggacttttctattaaaatataatgaaaatggaaccattacagatgctgtagactgtgcac
ttgaccctctctcagaaacaaagtgtacgttgaaatccttcactgtagaaaaaggaatctatcaaacttctaactttagagtcca
accaacagaatctattgttagatttcctaatattacaaacttgtgcccttttggtgaagtttttaacgccaccagatttgcatctgttt
atgcttggaacaggaagagaatcagcaactgtgttgctgattattctgtcctatataattccgcatcattttccacttttaagtgtt
atggagtgtctcctactaaattaaatgatctctgctttactaatgtctatgcagattcatttgtaattagaggtgatgaagtcagac
aaatcgctccagggcaaactggaaagattgctgattataattataaattaccagatgattttacaggctgcgttatagcttggaa
ttctaacaatcttgattctaaggttggtggtaattataattacctgtatagattgtttaggaagtctaatctcaaacctttgagaga
gatatttcaactgaaatctatcaggccggtagcacaccttgtaatggtgttgaaggttttaattgttactttcctttacaatcatatg
gtttccaacccactaatggtgttggttaccaaccatacagagtagtagtactttcttttgaacttctacatgcaccagcaactgttt
gtggacctaaaaagtctactaatttggttaaaaacaaatgtgtcaatttcaacttcaatggtttaacaggcacaggtgttcttact
gagtctaacaaaaagtttctgcctttccaacaatttggcagagacattgctgacactactgatgctgtccgtgatccacagaca
cttgagattcttgacattacaccatgttcttttggtggtgtcagtgttataacaccaggaacaaatacttctaaccaggttgctgtt
ctttatcagggtgttaactgcacagaagtccctgttgctattcatgcagatcaacttactcctacttggcgtgtttattctacaggt
tctaatgtttttcaaacacgtgcaggctgtttaataggggctgaacatgtcaacaactcatatgagtgtgacatacccattggtg
caggtatatgcgctagttatcagactcagactaattctcctcggcgggcacgtagtgtagctagtcaatccatcattgcctaca
ctatgtcacttggtgcagaaaattcagttgcttactctaataactctattgccatacccacaaattttactattagtgttaccacag
aaattctaccagtgtctatgaccaagacatcagtagattgtacaatgtacatttgtggtgattcaactgaatgcagcaatcttttg
ttgcaatatggcagtttttgtacacaattaaaccgtgctttaactggaatagctgttgaacaagacaaaaacacccaagaagttt
ttgcacaagtcaaacaaatttacaaaacaccaccaattaaagattttggtggttttaatttttcacaaatattaccagatccatcaa
aaccaagcaagaggtcatttattgaagatctacttttcaacaaagtgacacttgcagatgctggcttcatcaaacaatatggtg
attgccttggtgatattgctgctagagacctcatttgtgcacaaaagtttaacggccttactgttttgccacctttgctcacagat
gaaatgattgctcaatacacttctgcactgttagcgggtacaatcacttctggttggacctttggtgcaggtgctgcattacaaa
taccatttgctatgcaaatggcttataggtttaatggtattggagttacacagaatgttctctatgagaaccaaaaattgattgcc
aaccaatttaatagtgctattggcaaaattcaagactcactttcttccacagcaagtgcacttggaaaacttcaagatgtggtca
accaaaatgcacaagctttaaacacgcttgttaaacaacttagctccaattttggtgcaatttcaagtgttttaaatgatatccttt
cacgtcttgacaaagttgaggctgaagtgcaaattgataggttgatcacaggcagacttcaaagtttgcagacatatgtgact
caacaattaattagagctgcagaaatcagagcttctgctaatcttgctgctactaaaatgtcagagtgtgtacttggacaatca
aaaagagttgatttttgtggaaagggctatcatcttatgtccttccctcagtcagcacctcatggtgtagtcttcttgcatgtgact
tatgtccctgcacaagaaaagaacttcacaactgctcctgccatttgtcatgatggaaaagcacactttcctcgtgaaggtgtc
tttgtttcaaatggcacacactggtttgtaacacaaaggaatttttatgaaccacaaatcattactacagacaacacatttgtgtc

tggtaactgtgatgttgtaataggaattgtcaacaacacagtttatgatcctttgcaacctgaattagactcattcaaggaggag
ttagataaatattttaagaatcatacatcaccagatgttgatttaggtgacatctctggcattaatgcttcagttgtaaacattcaa
aaagaaattgaccgcctcaatgaggttgccaagaatttaaatgaatctctcatcgatctccaagaacttggaaagtatgagca
gtatataaaatggccatggtacatttggctaggttttatagctggcttgattgccatagtaatggtgacaattatgctttgctgtat
gaccagttgctgtagttgtctcaagggctgttgttcttgtggatcctgctgcaaatttgatgaagacgactctgagccagtgctc
aaaggagtcaaattacattacacataa

D614G Mutated Spike Protein Amino Acid Sequence SEQ ID NO:11

MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQD
LFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFG
TTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRV
YSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINL
VRDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAY
YVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNF
RVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNS
ASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYK
LPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGS
TPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKK
STNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQT
LEILDITPCSFGGVSVITPGTNTSNQVAVLYQGVNCTEVPVAIHADQLTPTWR
VYSTGSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPRRARSV
ASQSIIAYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYI
CGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIK
DFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLI
CAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQM
AYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNQ
NAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTYV
TQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGV
VFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEP
QIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDL
GDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLG
FIAGLIAIVMVTIMLCCMTSCCSCLKGCCSCGSCCKFDEDDSEPVLKGVKLHY
T

Figure 6

SARS RBD mRNA clone# 2874,
SARS2 Receptor Binding Domain (RBD)-foldon mRNA modified with Cleancap analogue and N1-Methyl-Pseudouridine-5'-Triphosphate. Codon optimized for human SEQ ID No.12

2874

ATTAATACGACTCACTATAAGGAGGCCACCATGTTCGTGTTTCTGGTGCTG
CTGCCTCTGGTGTCCAGCCAGCGGGTGCAGCCCACCGAATCCATCGTGCG
GTTCCCCAATATCACCAATCTGTGCCCCTTCGGCGAGGTGTTCAATGCCAC
CAGATTCGCCTCTGTGTACGCCTGGAACCGGAAGCGGATCAGCAATTGCG
TGGCCGACTACTCCGTGCTGTACAACTCCGCCAGCTTCAGCACCTTCAAGT
GCTACGGCGTGTCCCCTACCAAGCTGAACGACCTGTGCTTCACAAACGTG
TACGCCGACAGCTTCGTGATCCGGGGAGATGAAGTGCGGCAGATTGCCCC
TGGACAGACAGGCAAGATCGCCGACTACAACTACAAGCTGCCCGACGACT
TCACCGGCTGTGTGATTGCCTGGAACAGCAACAACCTGGACTCCAAAGTC
GGCGGCAACTACAATTACCTGTACCGGCTGTTCCGGAAGTCCAATCTGAA
GCCCTTCGAGCGGGACATCTCCACCGAGATCTATCAGGCCGGCAGCACCC
CTTGTAACGGCGTGGAAGGCTTCAACTGCTACTTCCCACTGCAGTCCTACG
GCTTTCAGCCCACAAATGGCGTGGGCTATCAGCCCTACAGAGTGGTGGTG
CTGAGCTTCGAACTGCTGCATGCCCCTGCCACAGTGTGCGGCCCTAAGAA
AAGCACCAATCTCGTGAAGAACAAATGCGTGAACTTCGGATCCGGAGGTG
GATACATCCCGGAGGCCCCTAGGGACGGTCAAGCTTACGTGAGAAAGGA
CGGCGAATGGGTTCTGCTGTCGACCTTCTTGGGATAAGCAAGCTGCTTAC
ATAGAACTCGCGGCGATTGGCATGCCGCCTTAAAATTTTTATTTTATTTTTT
CTTTTCTTTTCCGAATCGGATTTTGTTTTTAATATTTCAAAAAAAAAAAAA
AAAAAAAAAAAACGCGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAATTAATAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAATTAATAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAA AAAAAAAAAAAAAAAAA

Protein sequence for D614G Mutated Spike Protein ORF SEQ ID NO:13

MFVFLVLLPLVSSQRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRI
SNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIA
PGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLK
PFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSF
ELLHAPATVCGPKKSTNLVKNKCVNFGSGGGYIPEAPRDGQAYVRKDGEWV
LLST- FLG

Figure 7

18 polypeptide epitopes for 2019-nCoV spike proteins

SEQ ID	start	stop	peptide
14	576	589	RDIADTTDAVRDPQ
15	416	429	VRQIAPGQTGKIAD
16	684	696	QTQTNSPRRARSV
17	258	270	LTPGDSSSGWTAG
18	1146	1157	VYDPLQPELDSF
19	814	825	ILPDPSKPSKRS
20	504	515	YGFQPTNGVGYQ
21	1166	1176	KNHTSPDVDLG
22	482	492	YQAGSTPCNGV
23	80	90	SGTNGTKRFDN
24	30	40	RTQLPPAYTNS
25	1265	1274	FDEDDSEPVL
26	797	806	IYKTPPIKDF
27	606	615	VITPGTNTSN
28	532	541	TVCGPKKSTN
29	1078	1086	PAQEKNFTT
30	781	789	VEQDKNTQE
31	448	456	NNLDSKVGG

SEQ ID NO:32 - 49

The mRNA sequences corresponding to 18 epitopes of 2019-nCoV spike protein

SEQ ID	start	stop	Nucleotide sequence
32	576	589	AGGGACATCGCCGACACCACCGACGCCGTGAGGGACCCCCAG
33	416	429	GTGAGGCAGATCGCCCCCGGCCAGACCGGCAAGATCGCCGAC
34	684	696	CAGACCCAGACCAACAGCCCCAGGAGGGCCAGGAGCGTG
35	258	270	CTGACCCCCGGCGACAGCAGCAGCGGCTGGACCGCCGGC
36	1146	1157	GTGTACGACCCCCTGCAGCCCGAGCTGGACAGCTTC
37	814	825	ATCCTGCCCGACCCCAGCAAGCCCAGCAAGAGGAGC
38	504	515	TACGGCTTCCAGCCCACCAACGGCGTGGGCTACCAG
39	1166	1176	AAGAACCACACCAGCCCCGACGTGGACCTGGGC
40	482	492	TACCAGGCCGGCAGCACCCCCTGCAACGGCGTG

41	80	90	AGCGGCACCAACGGCACCAAGAGGTTCGACAAC
42	30	40	AGGACCCAGCTGCCCCCCGCCTACACCAACAGC
43	1265	1274	TTCGACGAGGACGACAGCGAGCCCGTGCTG
44	797	806	ATCTACAAGACCCCCCCCATCAAGGACTTC
45	606	615	GTGATCACCCCCGGCACCAACACCAGCAAC
46	532	541	ACCGTGTGCGGCCCCAAGAAGAGCACCAAC
47	1078	1086	CCCGCCCAGGAGAAGAACTTCACCACC
48	781	789	GTGGAGCAGGACAAGAACACCCAGGAG
49	448	456	AACAACCTGGACAGCAAGGTGGGCGGC

SEQ ID NO: 50-55

Six unique insertions identified by aligning against other coronavirus recognized by the human immune system

SEQ ID	start	stop	peptide	nucleotide
50	153	167	YYHKNNKSWMESEFR	TACTACCACAAGAACAACAAGAGCTGGATGGAGAGCGAGTTCAGG
51	76	87	AIHVSGTNGTKR	GCCATCCACGTGAGCGGCACCAACGGCACCAAGAGG
52	252	266	ALHRSYLTPGDSSSG	GCCCTGCACAGGAGCTACCTGACCCCCGGCGACAGCAGCAGCGGC
53	447	457	NNLDSKVGGNY	AACAACCTGGACAGCAAGGTGGGCGGCAACTAC
54	479	495	TEIYQAGSTPCNGVEGF	ACCGAGATCTACCAGGCCGGCAGCACCCCCTGCAACGGCGTGGAGGGCTTC
55	683	692	QTQTNSPRRA	CAGACCCAGACCAACAGCCCCAGGAGGGCC

SEQ ID NO:56 - 66

The RNA sequence coded for various viral peptides serving as potential antigens for vaccination against 2019-nCoV viral infections.

SEQ ID	Start	Stop	Peptide Sequence	Nucleotide Sequence
56	606	615	VITPGTNTSN	GUGAUCACCCCCGGCACCAACACCAGCAAC
57	532	541	TVCGPKKSTN	ACCGUGUGCGGCCCCAAGAAGAGCACCAAC

58	1078	1086	PAQEKNFTT	CCCGCCCAGGAGAAGAACUUCACCACC
59	781	789	VEQDKNTQE	GUGGAGCAGGACAAGAACACCCAGGAG
60	448	456	NNLDSKVGG	AACAACCUGGACAGCAAGGUGGGCGGC
61	153	167	YYHKNNKSWMESEFR	UACUACCACAAGAACAACAAGAGCUGGAUGGAGAGCGAGUUCAGG
62	76	87	AIHVSGTNGTKR	GCCAUCCACGUGAGCGGCACCAACGGCACCAAGAGG
63	252	266	ALHRSYLTPGDSSSG	GCCCUGCACAGGAGCUACCUGACCCCCGGCGACAGCAGCAGCGGC
64	447	457	NNLDSKVGGNY	AACAACCUGGACAGCAAGGUGGGCGGCAACUAC
65	479	495	TEIYQAGSTPCNGVEGF	ACCGAGAUCUACCAGGCCGGCAGCACCCCCUGCAACGGCGUGGAGGGCUUC
66	683	692	QTQTNSPRRA	CAGACCCAGACCAACAGCCCCAGGAGGGCC

Figure 8

SEQ ID No. 67
D614G_S6P_DNA_sequence

atgtttgtttttcttgttttattgccactagtctctagtcagtgtgttaatcttacaaccagaactcaattacccccctgcatacactaat
tctttcacacgtggtgtttattaccctgacaaagttttcagatcctcagttttacattcaactcaggacttgttcttacctttcttttcc
aatgttacttggttccatgctatacatgtctctgggaccaatggtactaagaggtttgataaccctgtcctaccatttaatgatgg
tgtttattttgcttccactgagaagtctaacataataagaggctggatttttggtactactttagattcgaagacccagtccctact
tattgttaataacgctactaatgttgttattaaagtctgtgaatttcaattttgtaatgatccatttttgggtgtttattaccacaaaaa
caacaaaagttggatggaaagtgagttcagagtttattctagtgcgaataattgcacttttgaatatgtctctcagccttttcttat
ggaccttgaaggaaaacagggtaatttcaaaaatcttagggaatttgtgtttaagaatattgatggttatttaaaatatattctaa
gcacacgcctattaatttagtgcgtgatctccctcagggtttttcggctttagaaccattggtagatttgccaataggtattaaca
tcactaggtttcaaacttacttgctttacatagaagttatttgactcctggtgattcttcttcaggttggacagctggtgctgcag
cttattatgtgggttatcttcaacctaggacttttctattaaaatataatgaaaatggaaccattacagatgctgtagactgtgcac
ttgaccctctctcagaaacaaagtgtacgttgaaatccttcactgtagaaaaaggaatctatcaaacttctaactttagagtcca
accaacagaatctattgttagatttcctaatattacaaacttgtgcccttttggtgaagtttttaacgccaccagatttgcatctgttt
atgcttggaacaggaagagaatcagcaactgtgttgctgattattctgtcctatataattccgcatcattttccacttttaagtgtt
atggagtgtctcctactaaattaaatgatctctgctttactaatgtctatgcagattcatttgtaattagaggtgatgaagtcagac
aaatcgctccagggcaaactggaaagattgctgattataattataaattaccagatgattttacaggctgcgttatagcttggaa
ttctaacaatcttgattctaaggttggtggtaattataattacctgtatagattgtttaggaagtctaatctcaaacctttgagaga
gatatttcaactgaaatctatcaggccggtagcacaccttgtaatggtgttgaaggttttaattgttactttcctttacaatcatatg
gtttccaacccactaatggtgttggttaccaaccatacagagtagtagtactttcttttgaacttctacatgcaccagcaactgttt
gtggacctaaaaagtctactaatttggttaaaaacaaatgtgtcaatttcaacttcaatggtttaacaggcacaggtgttcttact
gagtctaacaaaaagtttctgcctttccaacaatttggcagagacattgctgacactactgatgctgtccgtgatccacagaca
cttgagattcttgacattacaccatgttcttttggtggtgtcagtgttataacaccaggaacaaatacttctaaccaggttgctgtt
ctttatcagggtgttaactgcacagaagtccctgttgctattcatgcagatcaacttactcctacttggcgtgtttattctacaggt
tctaatgtttttcaaacacgtgcaggctgtttaataggggctgaacatgtcaacaactcatatgagtgtgacatacccattggtg
caggtatatgcgctagttatcagactcagactaattctccttcgcgggcaggtagtgtagctagtcaatccatcattgcctaca
ctatgtcacttggtgcagaaaattcagttgcttactctaataactctattgccatacccacaaattttactattagtgttaccacag
aaattctaccagtgtctatgaccaagacatcagtagattgtacaatgtacatttgtggtgattcaactgaatgcagcaatcttttg
ttgcaatatggcagtttttgtacacaattaaaccgtgctttaactggaatagctgttgaacaagacaaaaacacccaagaagttt
ttgcacaagtcaaacaaatttacaaaacaccaccaattaaagattttggtggttttaatttttcacaaatattaccagatccatcaa
aaccaagcaagaggtcacctattgaagatctacttttcaacaaagtgacacttgcagatgctggcttcatcaaacaatatggt
gattgccttggtgatattgctgctagagacctcatttgtgcacaaaagtttaacggccttactgttttgccacctttgctcacaga
tgaaatgattgctcaatacacttctgcactgttagcgggtacaatcacttctggttggacctttggtgcaggtcctgcattacaa
ataccatttcctatgcaaatggcttataggtttaatggtattggagttacacagaatgttctctatgagaaccaaaaattgattgc
caaccaatttaatagtgctattggcaaaattcaagactcactttcttccacaccaagtgcacttggaaaacttcaagatgtggtc
aaccaaaatgcacaagctttaaacacgcttgttaaacaacttagctccaattttggtgcaatttcaagtgttttaaatgatatcctt
tcacgtcttgacccacctgaggctgaagtgcaaattgataggttgatcacaggcagacttcaaagtttgcagacatatgtgac
tcaacaattaattagagctgcagaaatcagagcttctgctaatcttgctgctactaaaatgtcagagtgtgtacttggacaatca
aaaagagttgatttttgtggaaagggctatcatcttatgtccttccctcagtcagcacctcatggtgtagtcttcttgcatgtgact
tatgtccctgcacaagaaaagaacttcacaactgctcctgccatttgtcatgatggaaaagcacactttcctcgtgaaggtgtc
tttgtttcaaatggcacacactggtttgtaacacaaaggaatttttatgaaccacaaatcattactacagacaacacatttgtgtc

tggtaactgtgatgttgtaataggaattgtcaacaacacagtttatgatcctttgcaacctgaattagactcattcaaggaggag
ttagataaatattttaagaatcatacatcaccagatgttgatttaggtgacatctctggcattaatgcttcagttgtaaacattcaa
aaagaaattgaccgcctcaatgaggttgccaagaatttaaatgaatctctcatcgatctccaagaacttggaaagtatgagca
gtatataaaatggccatggtacatttggctaggttttatagctggcttgattgccatagtaatggtgacaattatgctttgctgtat
gaccagttgctgtagttgtctcaagggctgttgttcttgtggatcctgctgcaaatttgatgaagacgactctgagccagtgctc
aaaggagtcaaattacattacacataa

D614G_S6P_amino_acid_sequence SEQ ID No. 68

MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQD
LFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFG
TTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRV
YSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINL
VRDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAY
YVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNF
RVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNS
ASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYK
LPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGS
TPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKK
STNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQT
LEILDITPCSFGGVSVITPGTNTSNQVAVLYQGVNCTEVPVAIHADQLTPTWR
VYSTGSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPSRAGSV
ASQSIIAYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYI
CGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIK
DFGGFNFSQILPDPSKPSKRSPIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLI
CAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGPALQIPFPMQMA
YRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTPSALGKLQDVVNQNA
QALNTLVKQLSSNFGAISSVLNDILSRLDPPEAEVQIDRLITGRLQSLQTYVTQ
QLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVF
LHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQII
TTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLG
DISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFI
AGLIAIVMVTIMLCCMTSCCSCLKGCCSCGSCCKFDEDDSEPVLKGVKLHYT

Figure 9

RBD_trimer_DNA_sequence SEQ ID No. 69

Gccaccatgttcgtgtttctggtgctgctgcctctggtgtccagccagcgggtgcagcccaccgaatccatcgtgcggttcc
ccaatatcaccaatctgtgccccttcggcgaggtgttcaatgccaccagattcgcctctgtgtacgcctggaaccggaagcg
gatcagcaattgcgtggccgactactccgtgctgtacaactccgccagcttcagcaccttcaagtgctacggcgtgtcccct
accaagctgaacgacctgtgcttcacaaacgtgtacgccgacagcttcgtgatccggggagatgaagtgcggcagattgc
ccctggacagacaggcaagatcgccgactacaactacaagctgcccgacgacttcaccggctgtgtgattgcctggaaca
gcaacaacctggactccaaagtcggcggcaactacaattacctgtaccggctgttccggaagtccaatctgaagcccttcg
agcgggacatctccaccgagatctatcaggccggcagcacccttgtaacggcgtggaaggcttcaactgctacttcccac
tgcagtcctacggctttcagcccacaaatggcgtgggctatcagccctacagagtggtggtgctgagcttcgaactgctgca
tgcccctgccacagtgtgcggccctaagaaaagcaccaatctcgtgaagaacaaatgcgtgaacttcggatccggaggtg
gatacatcccggaggcccctagggacggtcaagcttacgtgagaaaggacggcgaatgggttctgctgtcgaccttcttg
ggataa

RBD_trimer_amino_acid_sequence SEQ ID No. 70

ATMFVFLVLLPLVSSQRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNR
KRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVR
QIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKS
NLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVV
LSFELLHAPATVCGPKKSTNLVKNKCVNFGSGGGYIPEAPRDGQAYVRKDGE
WVLLSTFLG

Figure 10

VUI-202012_01_S6P_DNA_sequence SEQ ID No. 71

atgtttgtttttcttgttttattgccactagtctctagtcagtgtgttaatcttacaaccagaactcaattacccccctgcatacactaat
tctttcacacgtggtgtttattaccctgacaaagttttcagatcctcagttttacattcaactcaggacttgttcttacctttcttttcc
aatgttacttggttccatgctatacatgtctctgggaccaatggtactaagaggtttgataaccctgtcctaccatttaatgatgg
tgtttattttgcttccactgagaagtctaacataataagaggctggatttttggtactactttagattcgaagacccagtccctact
tattgttaataacgctactaatgttgttattaaagtctgtgaatttcaattttgtaatgatccatttttgggtgtttattaccacaaaaa
caacaaaagttggatggaaagtgagttcagagtttattctagtgcgaataattgcacttttgaatatgtctctcagccttttcttat
ggaccttgaaggaaaacagggtaatttcaaaaatcttagggaatttgtgtttaagaatattgatggttatttaaaatatattctaa
gcacacgcctattaatttagtgcgtgatctccctcagggtttttcggctttagaaccattggtagatttgccaataggtattaaca
tcactaggtttcaaactttacttgctttacatagaagttatttgactcctggtgattcttcttcaggttggacagctggtgctgcag
cttattatgtgggttatcttcaacctaggacttttctattaaaatataatgaaaatggaaccattacagatgctgtagactgtgcac
ttgaccctctctcagaaacaaagtgtacgttgaaatccttcactgtagaaaaaggaatctatcaaacttctaactttagagtcca
accaacagaatctattgttagatttcctaatattacaaacttgtgcccttttggtgaagtttttaacgccaccagatttgcatctgttt
atgcttggaacaggaagagaatcagcaactgtgttgctgattattctgtcctatataattccgcatcattttccacttttaagtgtt
atggagtgtctcctactaaattaaatgatctctgctttactaatgtctatgcagattcatttgtaattagaggtgatgaagtcagac
aaatcgctccagggcaaactggaaagattgctgattataattataaattaccagatgattttacaggctgcgttatagcttggaa
ttctaacaatcttgattctaaggttggtggtaattataattacctgtatagattgtttaggaagtctaatctcaaacctttgagaga
gatatttcaactgaaatctatcaggccggtagcacaccttgtaatggtgttgaaggttttaattgttactttcctttacaatcatatg
gtttccaacccacttatggtgttggttaccaaccatacagagtagtagtactttcttttgaacttctacatgcaccagcaactgttt
gtggacctaaaaagtctactaatttggttaaaaacaaatgtgtcaatttcaacttcaatggtttaacaggcacaggtgttcttact
gagtctaacaaaaagtttctgcctttccaacaatttggcagagacattgatgacactactgatgctgtccgtgatccacagaca
cttgagattcttgacattacaccatgttcttttggtggtgtcagtgttataacaccaggaacaaatacttctaaccaggttgctgtt
ctttatcagggtgttaactgcacagaagtccctgttgctattcatgcagatcaacttactcctacttggcgtgtttattctacaggt
tctaatgtttttcaaacacgtgcaggctgtttaataggggctgaacatgtcaacaactcatatgagtgtgacatacccattggtg
caggtatatgcgctagttatcagactcagactaattctcattcgcgggcaggtagtgtagctagtcaatccatcattgcctaca
ctatgtcacttggtgcagaaaattcagttgcttactctaataactctattgccatacccataaattttactattagtgttaccacaga
aattctaccagtgtctatgaccaagacatcagtagattgtacaatgtacatttgtggtgattcaactgaatgcagcaatcttttgtt
gcaatatggcagtttttgtacacaattaaaccgtgctttaactggaatagctgttgaacaagacaaaaacacccaagaagttttt
gcacaagtcaaacaaatttacaaaacaccaccaattaaagattttggtggttttaatttttcacaaatattaccagatccatcaaa
accaagcaagaggtcacctattgaagatctacttttcaacaaagtgacacttgcagatgctggcttcatcaaacaatatggtg
attgccttggtgatattgctgctagagacctcatttgtgcacaaaagtttaacggccttactgttttgccacctttgctcacagat
gaaatgattgctcaatacacttctgcactgttagcgggtacaatcacttctggttggacctttggtgcaggtcctgcattacaaa
taccatttcctatgcaaatggcttataggtttaatggtattggagttacacagaatgttctctatgagaaccaaaaattgattgcc
aaccaatttaatagtgctattggcaaaattcaagactcactttcttccacaccaagtgcacttggaaaacttcaagatgtggtca
accaaaatgcacaagctttaaacacgcttgttaaacaacttagctccaattttggtgcaatttcaagtgttttaaatgatatccttg
cacgtcttgacccacctgaggctgaagtgcaaattgataggttgatcacaggcagacttcaaagtttgcagacatatgtgact
caacaattaattagagctgcagaaatcagagcttctgctaatcttgctgctactaaaatgtcagagtgtgtacttggacaatca
aaaagagttgatttttgtggaaagggctatcatcttatgtccttccctcagtcagcacctcatggtgtagtcttcttgcatgtgact
tatgtccctgcacaagaaaagaacttcacaactgctcctgccatttgtcatgatggaaaagcacactttcctcgtgaaggtgtc
tttgtttcaaatggcacacactggtttgtaacacaaaggaatttttatgaaccacaaatcattactacacacaacacatttgtgtct

ggtaactgtgatgttgtaataggaattgtcaacaacacagtttatgatcctttgcaacctgaattagactcattcaaggaggagt
tagataaatattttaagaatcatacatcaccagatgttgatttaggtgacatctctggcattaatgcttcagttgtaaacattcaaa
aagaaattgaccgcctcaatgaggttgccaagaatttaaatgaatctctcatcgatctccaagaacttggaaagtatgagcag
tatataaaatggccatggtacatttggctaggttttatagctggcttgattgccatagtaatggtgacaattatgctttgctgtatg
accagttgctgtagttgtctcaagggctgttgttcttgtggatcctgctgcaaatttgatgaagacgactctgagccagtgctca
aaggagtcaaattacattacacataa

VUI-202012_01_S6P_amino_acid_sequence SEQ ID No. 72

MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQD
LFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFG
TTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRV
YSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINL
VRDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAY
YVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNF
RVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNS
ASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYK
LPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGS
TPCNGVEGFNCYFPLQSYGFQPTYGVGYQPYRVVVLSFELLHAPATVCGPKK
STNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIDDTTDAVRDPQT
LEILDITPCSFGGVSVITPGTNTSNQVAVLYQGVNCTEVPVAIHADQLTPTWR
VYSTGSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSHSRAGSV
ASQSIIAYTMSLGAENSVAYSNNSIAIPINFTISVTTEILPVSMTKTSVDCTMYIC
GDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKD
FGGFNFSQILPDPSKPSKRSPIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLIC
AQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGPALQIPFPMQMA
YRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTPSALGKLQDVVNQNA
QALNTLVKQLSSNFGAISSVLNDILARLDPPEAEVQIDRLITGRLQSLQTYVTQ
QLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVF
LHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQII
TTHNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLG
DISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFI
AGLIAIVMVTIMLCCMTSCCSCLKGCCSCGSCCKFDEDDSEPVLKGVKLHYT

VUI-202012_del_S6P_DNA_sequence SEQ ID No. 73

atgtttgtttttcttgttttattgccactagtctctagtcagtgtgttaatcttacaaccagaactcaattacccccctgcatacactaat
tctttcacacgtggtgtttattaccctgacaaagttttcagatcctcagttttacattcaactcaggacttgttcttacctttctttttcc
aatgttacttggttccatgctatatctgggaccaatggtactaagaggtttgataaccctgtcctaccatttaatgatggtgtttatt
ttgcttccactgagaagtctaacataataagaggctggatttttggtactactttagattcgaagacccagtccctacttattgtta
ataacgctactaatgttgttattaaagtctgtgaatttcaattttgtaatgatccatttttgggtgtttaccacaaaaacaacaaaag
ttggatggaaagtgagttcagagtttattctagtgcgaataattgcacttttgaatatgtctctcagccttttcttatggaccttgaa

ggaaaacagggtaatttcaaaaatcttagggaatttgtgtttaagaatattgatggttatttttaaaatatattctaagcacacgcct
attaatttagtgcgtgatctccctcagggtttttcggctttagaaccattggtagatttgccaataggtattaacatcactaggtttc
aaactttacttgctttacatagaagttatttgactcctggtgattcttcttcaggttggacagctggtgctgcagcttattatgtggg
ttatcttcaacctaggacttttctattaaaatataatgaaaatggaaccattacagatgctgtagactgtgcacttgaccctctctc
agaaacaaagtgtacgttgaaatccttcactgtagaaaaaggaatctatcaaacttctaactttagagtccaaccaacagaat
ctattgttagatttcctaatattacaaacttgtgccctttggtgaagtttttaacgccaccagatttgcatctgtttatgcttggaac
aggaagagaatcagcaactgtgttgctgattattctgtcctatataattccgcatcattttccacttttaagtgttatggagtgtctc
ctactaaattaaatgatctctgctttactaatgtctatgcagattcatttgtaattagaggtgatgaagtcagacaaatcgctcca
gggcaaactggaaagattgctgattataattataaattaccagatgattttacaggctgcgttatagcttggaattctaacaatct
tgattctaaggttggtggtaattataattacctgtatagattgtttaggaagtctaatctcaaacctttgagagagatatttcaact
gaaatctatcaggccggtagcacaccttgtaatggtgttgaaggttttaattgttactttcctttacaatcatatggtttccaaccc
acttatggtgttggttaccaaccatacagagtagtagtactttcttttgaacttctacatgcaccagcaactgtttgtggacctaa
aaagtctactaatttggttaaaaacaaatgtgtcaatttcaacttcaatggtttaacaggcacaggtgttcttactgagtctaaca
aaaagtttctgcctttccaacaatttggcagagacattgatgacactactgatgctgtccgtgatccacagacacttgagattct
tgacattacaccatgttcttttggtggtgtcagtgttataacaccaggaacaaatacttctaaccaggttgctgttctttatcagg
gtgttaactgcacagaagtccctgttgctattcatgcagatcaacttactcctacttggcgtgtttattctacaggttctaatgttttt
caaacacgtgcaggctgtttaataggggctgaacatgtcaacaactcatatgagtgtgacatacccattggtgcaggtatatg
cgctagttatcagactcagactaattctcattcgcgggcaggtagtgtagctagtcaatccatcattgcctacactatgtcactt
ggtgcagaaaattcagttgcttactctaataactctattgccatacccataaattttactattagtgttaccacagaaattctacca
gtgtctatgaccaagacatcagtagattgtacaatgtacatttgtggtgattcaactgaatgcagcaatcttttgttgcaatatgg
cagtttttgtacacaattaaaccgtgctttaactggaatagctgttgaacaagacaaaaacacccaagaagtttttgcacaagt
caaacaaatttacaaaacaccaccaattaaagattttggtggttttaatttttcacaaatattaccagatccatcaaaaccaagca
agaggtcacctattgaagatctacttttcaacaaagtgacacttgcagatgctggcttcatcaaacaatatggtgattgccttgg
tgatattgctgctagagacctcatttgtgcacaaaagtttaacggccttactgttttgccacctttgctcacagatgaaatgattg
ctcaatacacttctgcactgttagcgggtacaatcacttctggttggacctttggtgcaggtcctgcattacaaataccatttcct
atgcaaatggcttataggtttaatggtattggagttacacagaatgttctctatgagaaccaaaaattgattgccaaccaatttaa
tagtgctattggcaaaattcaagactcactttcttccacaccaagtgcacttggaaaacttcaagatgtggtcaaccaaaatgc
acaagctttaaacacgcttgttaaacaacttagctccaattttggtgcaatttcaagtgttttaaatgatatccttgcacgtcttga
cccacctgaggctgaagtgcaaattgataggttgatcacaggcagacttcaaagtttgcagacatatgtgactcaacaattaa
ttagagctgcagaaatcagagcttctgctaatcttgctgctactaaaatgtcagagtgtgtacttggacaatcaaaaagagttg
atttttgtggaaagggctatcatcttatgtccttccctcagtcagcacctcatggtgtagtcttcttgcatgtgacttatgtccctgc
acaagaaaagaacttcacaactgctcctgccatttgtcatgatggaaaagcacactttcctcgtgaaggtgtctttgtttcaaat
ggcacacactggtttgtaacacaaaggaatttttatgaaccacaaatcattactacacacaacacatttgtgtctggtaactgtg
atgttgtaataggaattgtcaacaacacagtttatgatcctttgcaacctgaattagactcattcaaggaggagttagataaatat
tttaagaatcatacatcaccagatgttgatttaggtgacatctctggcattaatgcttcagttgtaaacattcaaaaagaaattga
ccgcctcaatgaggttgccaagaatttaaatgaatctctcatcgatctccaagaacttggaaagtatgagcagtatataaaatg
gccatggtacatttggctaggttttatagctggcttgattgccatagtaatggtgacaattatgctttgctgtatgaccagttgctg
tagttgtctcaagggctgttgttcttgtggatcctgctgcaaatttgatgaagacgactctgagccagtgctcaaaggagtcaa
attacattacacataa

VUI-202012_del_S6P_amino_acid_sequence SEQ ID No. 74

MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQD
LFLPFFSNVTWFHAISGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTT
LDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLGVYHKNNKSWMESEFRVYSS
ANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRD
LPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVG
YLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQ
PTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFS
TFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPD
DFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPC
NGVEGFNCYFPLQSYGFQPTYGVGYQPYRVVVLSFELLHAPATVCGPKKSTN
LVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIDDTTDAVRDPQTLEIL
DITPCSFGGVSVITPGTNTSNQVAVLYQGVNCTEVPVAIHADQLTPTWRVYST
GSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSHSRAGSVASQSI
IAYTMSLGAENSVAYSNNSIAIPINFTISVTTEILPVSMTKTSVDCTMYICGDST
ECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGF
NFSQILPDPSKPSKRSPIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQK
FNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGPALQIPFPMQMAYRFN
GIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTPSALGKLQDVVNQNAQALN
TLVKQLSSNFGAISSVLNDILARLDPPEAEVQIDRLITGRLQSLQTYVTQQLIRA
AEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVT
YVPAQEKNFTTAPAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTHN
TFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGI
NASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGLI
AIVMVTIMLCCMTSCCSCLKGCCSCGSCCKFDEDDSEPVLKGVKLHYT

Figure 11

SEQ ID No.75 – 88

Polymer	Branch Sequence	SEQ ID No.
-	R= $KH_nKH_nKH_nKH_nK$–	75
-	R= $H_nKH_nKH_nKH_nKH_nK$–	76
-	R= $KH_nKH_nKH_nKH_nKH_n$–	77
-	R= $H_nKH_nKH_nKH_nKH_nKH_n$–	78
H2K4b	R_A= KHKHHKHHKHHKHHKHHKHK–	79
H3K4b	R_B= KHHHKHHHKHHHKHHHK–	80
H3K(+H)4b	R_C= KHHHKHHHKH<u>H</u>HHKHHHK–	81
H3k(+H)4b	R_D= kHHHkHHHkH<u>H</u>HHkHHHk–	82
H-H3K(+H)4b	R_E=_ <u>H</u>KHHHKHHHKH<u>H</u>HHKHHHK–	83
HH-H3K(+H)4b	R_F=_ <u>HH</u>KHHHKHHHKH<u>H</u>HHKHHHK–	84
H4K4b	R_G= KH<u>H</u>HHKH<u>H</u>HHKH<u>H</u>HHKH<u>H</u>HHK–	85
H3K(1+H)4b	R_H= KHHHKHHHKHHHKH<u>H</u>HHK–	86
H3K(3+H)4b	R_I= KHHHKH<u>H</u>HHKHHHKHHHK–	87
H3K(1,3+H)4b	R_J= KHHHKH<u>H</u>HHKHHHKH<u>H</u>HHK–	88

5

Figure 12

D614G S2P DNA Sequence SEQ ID No.89:

atgtttgtttttcttgttttattgccactagtctctagtcagtgtgttaatcttacaaccagaactcaattacccccctgcatacactaat
tctttcacacgtggtgtttattaccctgacaaagttttcagatcctcagttttacattcaactcaggacttgttcttacctttcttttcc
aatgttacttggttccatgctatacatgtctctgggaccaatggtactaagaggtttgataaccctgtcctaccatttaatgatgg
tgtttattttgcttccactgagaagtctaacataataagaggctggatttttggtactactttagattcgaagacccagtccctact
tattgttaataacgctactaatgttgttattaaagtctgtgaatttcaattttgtaatgatccatttttgggtgtttattaccacaaaaa
caacaaaagttggatggaaagtgagttcagagtttattctagtgcgaataattgcacttttgaatatgtctctcagccttttcttat
ggaccttgaaggaaaacagggtaatttcaaaaatcttagggaatttgtgtttaagaatattgatggttatttaaaatatattctaa
gcacacgcctattaatttagtgcgtgatctccctcagggtttttcggctttagaaccattggtagatttgccaataggtattaaca
tcactaggtttcaaacttttacttgctttacatagaagttatttgactcctggtgattcttcttcaggttggacagctggtgctgcag
cttattatgtgggttatcttcaacctaggacttttctattaaaatataatgaaaatggaaccattacagatgctgtagactgtgcac
ttgaccctctctcagaaacaaagtgtacgttgaaatccttcactgtagaaaaaggaatctatcaaacttctaactttagagtcca
accaacagaatctattgttagatttcctaatattacaaacttgtgccctttggtgaagtttttaacgccaccagatttgcatctgttt
atgcttggaacaggaagagaatcagcaactgtgttgctgattattctgtcctatataattccgcatcattttccacttttaagtgtt
atggagtgtctcctactaaattaaatgatctctgctttactaatgtctatgcagattcatttgtaattagaggtgatgaagtcagac
aaatcgctccagggcaaactggaaagattgctgattataattataaattaccagatgattttacaggctgcgttatagcttggaa
ttctaacaatcttgattctaaggttggtggtaattataattacctgtatagattgtttaggaagtctaatctcaaacctttgagaga
gatatttcaactgaaatctatcaggccggtagcacaccttgtaatggtgttgaaggttttaattgttactttcctttacaatcatatg
gtttccaacccactaatggtgttggttaccaaccatacagagtagtagtactttcttttgaacttctacatgcaccagcaactgttt
gtggacctaaaaagtctactaatttggttaaaaacaaatgtgtcaatttcaacttcaatggtttaacaggcacaggtgttcttact
gagtctaacaaaaagtttctgcctttccaacaatttggcagagacattgctgacactactgatgctgtccgtgatccacagaca
cttgagattcttgacattacaccatgttcttttggtggtgtcagtgttataacaccaggaacaaatacttctaaccaggttgctgtt
ctttatcagggtgttaactgcacagaagtccctgttgctattcatgcagatcaacttactcctacttggcgtgtttattctacaggt
tctaatgtttttcaaacacgtgcaggctgtttaataggggctgaacatgtcaacaactcatatgagtgtgacatacccattggtg
caggtatatgcgctagttatcagactcagactaattctccttcgcgggcaggtagtgtagctagtcaatccatcattgcctaca
ctatgtcacttggtgcagaaaattcagttgcttactctaataactctattgccatacccacaaattttactattagtgttaccacag
aaattctaccagtgtctatgaccaagacatcagtagattgtacaatgtacatttgtggtgattcaactgaatgcagcaatcttttg
ttgcaatatggcagtttttgtacacaattaaaccgtgctttaactggaatagctgttgaacaagacaaaaacacccaagaagttt
ttgcacaagtcaaacaaatttacaaaacaccaccaattaaagattttggtggttttaatttttcacaaatattaccagatccatcaa
aaccaagcaagaggtcatttattgaagatctacttttcaacaaagtgacacttgcagatgctggcttcatcaaacaatatggtg
attgccttggtgatattgctgctagagacctcatttgtgcacaaaagtttaacggccttactgttttgccacctttgctcacagat
gaaatgattgctcaatacacttctgcactgttagcgggtacaatcacttctggttggacctttggtgcaggtgctgcattacaaa
taccatttgctatgcaaatggcttataggtttaatggtattggagttacacagaatgttctctatgagaaccaaaaattgattgcc
aaccaatttaatagtgctattggcaaaattcaagactcactttcttccacagcaagtgcacttggaaaacttcaagatgtggtca
accaaaatgcacaagctttaaacacgcttgttaaacaacttagctccaattttggtgcaatttcaagtgttttaaatgatatccttt
cacgtcttgacccacctgaggctgaagtgcaaattgataggttgatcacaggcagacttcaaagtttgcagacatatgtgact
caacaattaattagagctgcagaaatcagagcttctgctaatcttgctgctactaaaatgtcagagtgtgtacttggacaatca
aaaagagttgatttttgtggaaagggctatcatcttatgtccttccctcagtcagcacctcatggtgtagtcttcttgcatgtgact
tatgtccctgcacaagaaaagaacttcacaactgctcctgccatttgtcatgatggaaaagcacactttcctcgtgaaggtgtc
tttgtttcaaatggcacacactggtttgtaacacaaaggaatttttatgaaccacaaatcattactacagacaacacatttgtgtc
tggtaactgtgatgttgtaataggaattgtcaacaacacagtttatgatcctttgcaacctgaattagactcattcaaggaggag

ttagataaatattttaagaatcatacatcaccagatgttgatttaggtgacatctctggcattaatgcttcagttgtaaacattcaa
aaagaaattgaccgcctcaatgaggttgccaagaatttaaatgaatctctcatcgatctccaagaacttggaaagtatgagca
gtatataaaatggccatggtacatttggctaggttttatagctggcttgattgccatagtaatggtgacaattatgctttgctgtat
gaccagttgctgtagttgtctcaagggctgttgttcttgtggatcctgctgcaaatttgatgaagacgactctgagccagtgctc
aaaggagtcaaattacattacacataa

Figure 13

D614G S2P amino acid sequence SEQ ID No.90

MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQD
LFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFG
TTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRV
YSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINL
VRDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAY
YVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNF
RVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNS
ASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYK
LPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGS
TPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKK
STNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQT
LEILDITPCSFGGVSVITPGTNTSNQVAVLYQGVNCTEVPVAIHADQLTPTWR
VYSTGSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPSRAGSV
ASQSIIAYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYI
CGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIK
DFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLI
CAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQM
AYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNQ
NAQALNTLVKQLSSNFGAISSVLNDILSRLDPPEAEVQIDRLITGRLQSLQTYV
TQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGV
VFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEP
QIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDL
GDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLG
FIAGLIAIVMVTIMLCCMTSCCSCLKGCCSCGSCCKFDEDDSEPVLKGVKLHY
T

Figure 14

SEQ ID No.90 to 95

SEQ ID	start	stop	nucleotide
90	153	167	TACTACCACAAGAACAACAAGAGCTGGATG GAGAGCGAGTTCAGG
91	76	87	GCCATCCACGTGAGCGGCACCAACGGCACC AAGAGG
92	252	266	GCCCTGCACAGGAGCTACCTGACCCCCGGC GACAGCAGCAGCGGC
93	447	457	AACAACCTGGACAGCAAGGTGGGCGGCAAC TAC
94	479	495	ACCGAGATCTACCAGGCCGGCAGCACCCCC TGCAACGGCGTGGAGGGCTTC
95	683	692	CAGACCCAGACCAACAGCCCCAGGAGGGCC

SEQ ID No. 96 to 106

SEQ ID	Start	Stop	Nucleotide Sequence
96	606	615	GUGAUCACCCCCGGCACCAACACCAGCAA C
97	532	541	ACCGUGUGCGGCCCCAAGAAGAGCACCAA C
98	1078	1086	CCCGCCCAGGAGAAGAACUUCACCACC
99	781	789	GUGGAGCAGGACAAGAACACCCAGGAG
100	448	456	AACAACCUGGACAGCAAGGUGGGCGGC
101	153	167	UACUACCACAAGAACAACAAGAGCUGGAU GGAGAGCGAGUUCAGG
102	76	87	GCCAUCCACGUGAGCGGCACCAACGGCAC CAAGAGG
103	252	266	GCCCUGCACAGGAGCUACCUGACCCCCGG CGACAGCAGCAGCGGC
104	447	457	AACAACCUGGACAGCAAGGUGGGCGGCAA CUAC

105	479	495	ACCGAGAUCUACCAGGCCGGCAGCACCCC CUGCAACGGCGUGGAGGGCUUC
106	683	692	CAGACCCAGACCAACAGCCCCAGGAGGGC C