Innovation, Sciences et Développement économique Canada
Office de la Propriété Intellectuelle du Canada

Innovation, Science and Economic Development Canada
Canadian Intellectual Property Office

CA 3215771 A1 2022/10/27

(21) 3 215 771

(12) DEMANDE DE BREVET CANADIEN
CANADIAN PATENT APPLICATION

(13) A1

(86) **Date de dépôt PCT/PCT Filing Date:** 2022/04/20
(87) **Date publication PCT/PCT Publication Date:** 2022/10/27
(85) **Entrée phase nationale/National Entry:** 2023/10/17
(86) **N° demande PCT/PCT Application No.:** EP 2022/060417
(87) **N° publication PCT/PCT Publication No.:** 2022/223617
(30) **Priorité/Priority:** 2021/04/20 (EP PCT/EP2021/060273)

(51) **Cl.Int./Int.Cl.** ***A61K 39/12*** (2006.01),
A61P 31/14 (2006.01), ***C12N 15/09*** (2006.01)

(71) **Demandeur/Applicant:**
BIONTECH SE, DE

(72) **Inventeurs/Inventors:**
SAHIN, UGUR, DE;
MUIK, ALEXANDER, DE

(74) **Agent:** BERESKIN & PARR LLP/S.E.N.C.R.L.,S.R.L.

(54) **Titre : VACCIN ANTIVIRAL**
(54) **Title: VIRUS VACCINE**

(57) **Abrégé/Abstract:**
This disclosure relates to the field of preventing or treating virus infection, in particular, the disclosure relates to agents for vaccination against virus infection and inducing effective virus antigen-specific immune responses such as antibody and/or T cell responses and methods for generating and using such agents. Administration of agents such as RNA disclosed herein to a subject can protect the subject against virus infection. Specifically, the present disclosure relates to amino acid sequences comprising at least a portion of a virus protein having amino acid modifications found in other variants of the virus protein. Administration of RNA encoding one or more of the amino acid sequences may provide protection against diverse virus variants. Methods and agents described herein are, in particular, useful for the prevention or treatment of coronavirus infection such as SARS-CoV-2 infection.

Date Submitted: 2023/10/17

CA App. No.: 3215771

Abstract:

This disclosure relates to the field of preventing or treating virus infection, in particular, the disclosure relates to agents for vaccination against virus infection and inducing effective virus antigen-specific immune responses such as antibody and/or T cell responses and methods for generating and using such agents. Administration of agents such as RNA disclosed herein to a subject can protect the subject against virus infection. Specifically, the present disclosure relates to amino acid sequences comprising at least a portion of a virus protein having amino acid modifications found in other variants of the virus protein. Administration of RNA encoding one or more of the amino acid sequences may provide protection against diverse virus variants. Methods and agents described herein are, in particular, useful for the prevention or treatment of coronavirus infection such as SARS-CoV-2 infection.

VIRUS VACCINE

This disclosure relates to the field of preventing or treating virus infection. In particular, the disclosure relates to agents for vaccination against virus infection and inducing effective virus antigen-specific immune responses such as antibody and/or T cell responses and methods for generating and using such agents. Administration of agents such as RNA disclosed herein to a subject can protect the subject against virus infection. Specifically, the present disclosure relates to amino acid sequences comprising at least a portion of a virus protein having amino acid modifications found in other variants of the virus protein. Administration of RNA encoding one or more of the amino acid sequences may provide protection against diverse virus variants. Methods and agents described herein are, in particular, useful for the prevention or treatment of coronavirus infection such as SARS-CoV-2 infection.

As a virus replicates, its genes undergo random "copying errors" (i.e. genetic mutations). Over time, these genetic copying errors can, among other changes to the virus, lead to alterations in the surface proteins or antigens of the virus. Genetic mutations may cause virus antigens to "drift"- meaning the surface of the mutated virus looks different than the original virus. When a virus drifts enough, vaccines against old strains of the virus and immunity from previous virus infections may no longer be effective against the new, drifted strains. A person then may become vulnerable to the newer, mutated viruses.

Coronaviruses are positive-sense, single-stranded RNA ((+)ssRNA) enveloped viruses that encode for a total of four structural proteins, spike protein (S), envelope protein (E), membrane protein (M) and nucleocapsid protein (N). The spike protein (S protein) is responsible for receptor-recognition, attachment to the cell, infection via the endosomal pathway, and the genomic release driven by fusion of viral and endosomal membranes. Though sequences between the different family members vary, there are conserved regions and motifs within the S protein making it possible to divide the S protein into two subdomains: S1 and S2. While the S2, with its transmembrane domain, is responsible for membrane fusion, the S1 domain recognizes the virus-specific receptor and binds to the target host cell. Within

CA 03215771 2023-10-17

several coronavirus isolates, the receptor binding domain (RBD) was identified and a general structure of the S protein defined (Figure 1).

In December 2019, a pneumonia outbreak of unknown cause occurred in Wuhan, China and it became clear that a novel coronavirus (severe acute respiratory syndrome coronavirus 2; SARS-CoV-2) was the underlying cause. The genetic sequence of SARS-CoV-2 became available to the WHO and public (MN908947.3) and the virus was categorized into the betacoronavirus subfamily. By sequence analysis, the phylogenetic tree revealed a closer relationship to severe acute respiratory syndrome (SARS) virus isolates than to another coronavirus infecting humans, namely the Middle East respiratory syndrome (MERS) virus.

SARS-CoV-2 infections and the resulting disease COVID-19 have spread globally, affecting a growing number of countries. On 11 March 2020 the WHO characterized the COVID-19 outbreak as a pandemic. As of 01 December 2020, there have been >63 million globally confirmed COVID-19 cases and >1.4 million deaths, with 191 countries/regions affected. The ongoing pandemic remains a significant challenge to public health and economic stability worldwide.

Every individual is at risk of infection as there is no pre-existing immunity to SARS-CoV-2. Following infection some but not all individuals develop protective immunity in terms of neutralising antibody responses and cell mediated immunity. However, it is currently unknown to what extent and for how long this protection lasts. According to WHO 80% of infected individuals recover without need for hospital care, while 15% develop more severe disease and 5% need intensive care. Increasing age and underlying medical conditions are considered risk factors for developing severe disease.

The presentation of COVID-19 is generally with cough and fever, with chest radiography showing ground-glass opacities or patchy shadowing. However, many patients present without fever or radiographic changes, and infections may be asymptomatic which is relevant to controlling transmission. For symptomatic subjects, progression of disease may lead to acute respiratory distress syndrome requiring ventilation and subsequent multi-organ failure and death. Common symptoms in hospitalized patients (in order of highest to lowest frequency) include fever, dry cough, shortness of breath, fatigue, myalgias, nausea/vomiting or diarrhoea, headache, weakness, and rhinorrhoea. Anosmia (loss of smell) or ageusia (loss

CA 03215771 2023-10-17

of taste) may be the sole presenting symptom in approximately 3% of individuals who have COVID-19.

All ages may present with the disease, but notably case fatality rates (CFR) are elevated in persons >60 years of age. Comorbidities are also associated with increased CFR, including cardiovascular disease, diabetes, hypertension, and chronic respiratory disease. Healthcare workers are overrepresented among COVID-19 patients due to occupational exposure to infected patients.

In most situations, a molecular test is used to detect SARS-CoV-2 and confirm infection. The reverse transcription polymerase chain reaction (RT-PCR) test methods targeting SARS-CoV-2 viral RNA are the gold standard in vitro methods for diagnosing suspected cases of COVID-19. Samples to be tested are collected from the nose and/or throat with a swab.

SARS-CoV-2 is an RNA virus with four structural proteins. One of them, the spike protein is a surface protein which binds the angiotensin-converting enzyme 2 (ACE-2) present on host cells. Therefore, the spike protein is considered a relevant antigen for vaccine development.

BNT162b2 is an mRNA vaccine for prevention of COVID-19 and demonstrated an efficacy of 95% or more at preventing COVID-19. The vaccine is made of a 5'capped mRNA encoding for the full-length SARS-CoV-2 spike glycoprotein (S) encapsulated in lipid nanoparticles (LNPs). The finished product is presented as a concentrate for dispersion for injection containing BNT162b2 as active substance. Other ingredients are: ALC-0315 (4-hydroxybutyl)azanediyl)bis(hexane-6,1-diyl)bis(2-hexyldecanoate), ALC-0159 (2-[(polyethylene glycol)-2000]-N,N-ditetradecylacetamide), 1,2-Distearoyl-sn-glycero-3-phosphocholine (DSPC), cholesterol, potassium chloride, potassium dihydrogen phosphate, sodium chloride, disodium phosphate dihydrate, sucrose and water for injection.

The sequence of the S protein was chosen based on the sequence for the "SARS-CoV-2 isolate Wuhan-Hu -1", which was available when the program was initiated: GenBank: MN908947.3 (complete genome) and GenBank: QHD43416.1 (spike surface glycoprotein). The active substance consists of a single-stranded, 5'-capped codon-optimized mRNA that is translated into the spike antigen of SARS-CoV-2. The protein sequence contains two proline mutations, which ensure an antigenically optimal pre-fusion confirmation (P2 S). The RNA does not contain any uridines; instead of uridine the modified N1-methylpseudouridine is used in RNA

CA 03215771 2023-10-17

synthesis. The RNA contains common structural elements optimized for mediating high RNA stability and translational efficiency. The LNPs protect the RNA from degradation by RNAses and enable transfection of host cells after intramuscular (IM) delivery. The mRNA is translated into the SARS-CoV-2 S protein in the host cell. The S protein is then expressed on the cell surface where it induces an adaptive immune response. The S protein is identified as a target for neutralising antibodies against the virus and is therefore considered a relevant vaccine component. BNT162b2 is administered intramuscularly (IM) in two 30 µg doses of the diluted vaccine solution given 21 days apart.

The recent emergence of novel circulating variants of SARS-CoV-2 has raised significant concerns about geographic and temporal efficacy of vaccine interventions. One of the earliest variants that emerged and rapidly became globally dominant was D614G. Furthermore, recent genomic surveillance in the United Kingdom has revealed rapid expansion of a novel lineage termed B.1.1.7 (also known as VOC-202012/01 or 501Y.V1). B.1.1.7 harbors deletions of three amino acids and seven missense mutations in spike, including D614G as well as N501Y in the ACE2 receptor-binding domain (RBD). It has been shown to be inherently more transmissible, with a growth rate that has been estimated to be 40-70% higher than other SARS-CoV-2 lineages in multiple countries (Volz et al., 2021, Nature, https://doi.org/10.1038/s41586-021-03470-x; Washington et al., 2021, Cell https://doi.org/10.1016/j.cell.2021.03.052).

Studies have demonstrated that BNT162b2 vaccine-elicited human sera cross-neutralize B.1.1.7 variants, suggesting that prior infection or vaccination with wild-type SARS-CoV-2 may still provide protection against B.1.1.7 variants (Muik A. et al., 2021, Science 371(6534):1152-1153).

There have also been reports of SARS-CoV-2 transmission between humans and minks in Denmark with a variant called mink cluster 5 or B.1.1.298, which harbors a deletion of two amino acids and four missense mutations including Y453F in RBD.

Other variants that recently emerged in California, United States, designated as B.1.427/B.1.429, contain four missense mutations in spike, one of which is a single L452R RBD mutation.

Novel variants arising from the B.1.1.28 lineage first described in Brazil and Japan, termed P.2 (with 3 spike missense mutations) and P.1 (with 12 spike missense mutations), contain an

CA 03215771 2023-10-17

E484K mutation in RBD, which is of particular concern, and P.1 in particular also contains K417T and N501Y mutations in RBD.

The emergence of multiple strains of the B.1.351 lineage (also known as 501Y.V2), which were first reported in South Africa and have since spread globally, are of great concern. This lineage bears three RBD mutations, K417N, E484K, and N501Y, in addition to several mutations outside of RBD. Using BNT162b2-elicited serum it was reported that, as compared with neutralization of the USA-WA1/2020 strain, neutralization of B.1.1.7-spike and P.1-spike viruses was roughly equivalent, and neutralization of B.1.351-spike virus was robust but lower (Liu Y. et al., 2021, N Engl J Med., doi: 10.1056/NEJMc2102017. Epub ahead of print. PMID: 33684280). In view of the emergence of novel variants that appear to at least partially escape immune responses there is a need for vaccines that are effective against variants of SARS-CoV-2.

The present disclosure provides vaccines that include epitopes of diverse virus protein variants and methods for providing such vaccines. Instead of re-formulating existing vaccines to include diverse virus protein sequences of different virus variants or the coding nucleic acids, the strategy described herein is based on combining diverse epitopes of different virus variants on a limited number of molecules. This allows for the administration of a relatively low dose of the active ingredient, in particular mRNA, while achieving sufficiently high doses of each epitope to induce an effective immune response. It is envisioned that the vaccines described herein are capable of eliciting broadly neutralizing antibodies and thus, are suitable to resolve the ongoing SARS-CoV-2 pandemic.

Summary

The present invention generally relates to a vaccine wherein epitopes of different variants of a virus protein are combined on a single molecule. Such molecule comprises at least a portion of a virus protein comprising amino acid modifications present in other variants of the virus protein. The modification generates (additional) epitopes that are specific for such other virus protein variants. Accordingly, the modified virus protein sequences described herein are polyspecific virus protein amino acid sequences. In one embodiment, a polyspecific virus

protein amino acid sequence is a modified full-length virus protein. In one embodiment, a polyspecific virus protein amino acid sequence is a modified portion of a virus protein.

In some embodiments, a polyspecific virus protein amino acid sequence provided herein may comprise immunoreactive epitopes ((modified) amino acid residues) from a plurality (e.g., at least two or more, including, e.g., at least three, at least four, at least five, at least six, at least seven, at least eight, at least nine, at least ten, etc.) of virus protein variants, e.g., immunoreactive epitopes from the parental virus protein and additionally immunoreactive epitopes from one or more virus protein variants which are different from the parental virus protein. In various embodiments, a polyspecific virus protein amino acid sequence provided herein may comprise all immunoreactive epitopes ((modified) amino acid residues) from one or more virus protein variants which are different from the parental virus protein, or a portion thereof. In various embodiments, a polyspecific virus protein amino acid sequence provided herein may comprise a plurality (e.g., at least two or more, including, e.g., at least three, at least four, at least five, at least six, at least seven, at least eight, at least nine, at least ten, etc.) of immunoreactive epitopes ((modified) amino acid residues) from one or more virus protein variants which are different from the parental virus protein.

In some embodiments, an amino acid sequence comprising a polyspecific virus protein amino acid sequence is encoded by a nucleic acid such as DNA or RNA, in particular RNA. In some embodiments, a plurality of amino acid sequences each of which comprises a polyspecific virus protein amino acid sequence is encoded by a nucleic acid such as DNA or RNA, in particular RNA. In some embodiments, a plurality of amino acid sequences each of which comprises a polyspecific virus protein amino acid sequence is encoded by a plurality of nucleic acid molecules such as DNA or RNA molecules, in particular RNA molecules.

The amino acid sequences comprising polyspecific virus protein amino acid sequences described herein and nucleic acids encoding these amino acid sequences are particularly useful, when considering the genetic diversity of viruses such as RNA viruses, to provide protection against a plurality of virus variants. The amino acid sequences comprising polyspecific virus protein amino acid sequences described herein and nucleic acids encoding these amino acid sequences may offer an opportunity for development of a diverse and/or otherwise robust (e.g., persistent, e.g., detectable about 5, 10, 15, 20, 25, 30, 35, 40, 45, 50,

CA 03215771 2023-10-17

55, 60 or more days after administration of one or more doses) neutralizing antibody and/or T cell response, e.g., TH1-type T cell (e.g., CD4+ and/or CD8+ T cell) response. The amino acid sequences comprising polyspecific virus protein amino acid sequences described herein and nucleic acids encoding the amino acid sequences are expected to elicit immune responses, in particular antibody response, that broadly and specifically neutralize a plurality (e.g., at least two or more, including, e.g., at least three, at least four, at least five, at least six, at least seven, at least eight, at least nine, at least ten, etc.) of virus variant strains and have the potential to generate protective immune responses to a range of virus variant strains in a limited number of constructs. In general, the limited number of constructs includes a number of molecules (protein and/or nucleic acid molecules) which is less than the number of the different virus protein variants, e.g., 1, 2, 3, or 4 molecules, such that each molecule includes epitope sequences that correspond to a plurality of virus protein variants. In general, the amino acid sequences comprising polyspecific virus protein amino acid sequences described herein , i.e., vaccine antigens, and nucleic acids encoding these amino acid sequences are useful for the immunotherapeutic treatment of a subject. A vaccine antigen comprises virus protein epitopes that are derived from and specific for inducing an immune response against a plurality of virus protein variants, and thus, virus variant strains, in the subject. In one embodiment, the present invention comprises the administration of nucleic acid such as RNA, i.e., vaccine RNA, encoding one or more of the vaccine antigens described herein. In one embodiment, the present invention comprises the administration of a plurality, e.g., 2, 3, or 4, of nucleic acid molecules such as RNA molecules encoding different vaccine antigens. The different vaccine antigens (potentially based on the same parental virus protein sequence) may comprise different modifications (optionally from different virus strains) and thus, different immunogenic spectra. RNA encoding vaccine antigen may be administered to provide (following expression of the RNA by appropriate target cells) antigen for induction, i.e., stimulation, priming and/or expansion, of an immune response, e.g., antibodies and/or immune effector cells, which is targeted to target antigen (virus protein, in particular different virus protein variants) or a procession product thereof. In one embodiment, the immune response which is to be induced according to the present disclosure is a B cell-mediated immune response, i.e., an antibody-mediated immune response. Additionally or alternatively,

CA 03215771 2023-10-17

in one embodiment, the immune response which is to be induced according to the present disclosure is a T cell-mediated immune response. In one embodiment, the immune response is an anti-virus immune response. In one embodiment, the immune response is an immune response which is directed against a plurality of virus strains.

In one aspect, the invention relates to a method comprising the steps:
a) identifying amino acid positions in a parental virus protein which are modified compared to the corresponding amino acid positions of one or more virus protein variants; and
b) providing an amino acid sequence comprising at least a fragment of the parental virus protein, wherein amino acid positions in the at least a fragment of a parental virus protein are modified to comprise amino acids found in the corresponding amino acid positions of one or more virus protein variants, or a nucleotide sequence encoding the modified amino acid sequence.
In one embodiment, the method comprises repeating step b) to provide two or more of the modified amino acid sequences, or two or more of the nucleotide sequences encoding two or more of the modified amino acid sequences.
In one embodiment, the two or more modified amino acid sequences are based on the same parental virus protein.
In one embodiment, the amino acid modifications in the two or more modified amino acid sequences are at least partially different.
In one embodiment, providing the nucleotide sequence comprises:
b') substituting codons of a nucleotide sequence encoding an amino acid sequence comprising at least a fragment of a parental virus protein with other codons to obtain a mutated nucleotide sequence that encodes a modified amino acid sequence, wherein amino acid positions in the at least a fragment of a parental virus protein are modified to comprise amino acids found in the corresponding amino acid positions of one or more virus protein variants.
In one embodiment, the method comprises repeating step b') to provide two or more of the mutated nucleotide sequences encoding two or more of the modified amino acid sequences.
In one embodiment, the two or more modified amino acid sequences are based on the same parental virus protein.

In one embodiment, the amino acid modifications in the two or more modified amino acid sequences are at least partially different.

In one aspect, the invention relates to a method comprising the steps:

a) identifying amino acid positions in a parental virus protein which are modified compared to the corresponding amino acid positions of one or more virus protein variants;

b) substituting codons of a first nucleotide sequence encoding an amino acid sequence comprising at least a fragment of a parental virus protein with other codons to obtain a first mutated nucleotide sequence that encodes a modified amino acid sequence, wherein amino acid positions in the at least a fragment of a parental virus protein are modified to comprise amino acids found in the corresponding amino acid positions of one or more virus protein variants; and

c) substituting codons of a second nucleotide sequence encoding an amino acid sequence comprising at least a fragment of a parental virus protein with other codons to obtain a second mutated nucleotide sequence that encodes a modified amino acid sequence, wherein amino acid positions in the at least a fragment of a parental virus protein are modified to comprise amino acids found in the corresponding amino acid positions of one or more virus protein variants,

wherein the amino acid modifications in b) at least partially differ from the amino acid modifications in c).

In one embodiment, the method may comprise one or more further such steps of substituting codons to obtain a mutated nucleotide sequence encoding at least partially different amino acid modifications.

In one embodiment, the at least a fragment of a parental virus protein comprised in the amino acid sequence encoded by the first nucleotide sequence and the at least a fragment of a parental virus protein comprised in the amino acid sequence encoded by the second nucleotide sequence are identical.

In one embodiment, the amino acid sequence encoded by the first nucleotide sequence and the amino acid sequence encoded by the second nucleotide sequence are identical.

In one embodiment, the first nucleotide sequence and the second nucleotide sequence are identical.

In one embodiment, one or more of the modified amino acid positions in the modified amino acid sequence encoded by the first mutated nucleotide sequence differ from the modified amino acid positions in the modified amino acid sequence encoded by the second mutated nucleotide sequence, i.e., one or more different positions are modified.

In one embodiment, one or more amino acids in modified amino acid positions modified in the modified amino acid sequence encoded by the first mutated nucleotide sequence and in the modified amino acid sequence encoded by the second mutated nucleotide sequence differ from each other, i.e., one or more of the same positions are modified, however, with different amino acid residues.

In one embodiment, the amino acid sequence comprising at least a fragment of a parental virus protein comprises the amino acid sequence of a full-length virus protein.

In one embodiment, the method further comprises providing nucleic acid comprising the nucleotide sequence encoding a modified amino acid sequence.

In one embodiment, the method further comprises providing a vaccine comprising nucleic acid comprising the nucleotide sequence encoding a modified amino acid sequence.

In one embodiment, the nucleic acid is RNA.

In one embodiment, the method is a method of generating a vaccine.

In one embodiment, the vaccine is an RNA vaccine.

In one embodiment, the vaccine has a reduced risk for immune escape.

In one embodiment, the modified amino acid positions are amino acid positions at which the amino acid sequence of the one or more virus protein variants differs from the amino acid sequence of the parental virus protein.

In one embodiment, the modified amino acid positions are amino acid positions at which the amino acid sequence of the one or more virus protein variants differs from the amino acid sequence of the wildtype virus protein.

In one embodiment, the modified amino acid positions are potential sites for escape mutants of the virus.

In one embodiment, the escape mutants of the virus are antibody escape mutants of the virus.

In one embodiment, the escape mutants of the virus are resistant to neutralization by antibody against the virus protein.

CA 03215771 2023-10-17

In one embodiment, the virus protein of the escape mutants of the virus shows reduced antibody binding.

In one embodiment, the antibody is used for treating a patient infected with the virus.

In one embodiment, the antibody is generated in a patient that has been treated with a vaccine against infection with the virus.

In one embodiment, the parental virus protein is modified compared to the wildtype virus protein.

In one embodiment, in the modified amino acid sequence amino acid positions in the parental virus protein which are modified compared to the wildtype virus protein are not modified.

In one embodiment, the parental virus protein is the virus protein of a parental virus strain.

In one embodiment, the parental virus strain is a natural isolate, or the parental virus strain is a mutant of a natural isolate.

In one embodiment, the parental virus strain is a virus variant strain that is prevalent or rapidly spreading.

In one embodiment, the parental virus strain is a virus variant that is a variant of concern.

In one embodiment, the one or more virus protein variants are modified compared to the wildtype virus protein.

In one embodiment, the one or more virus protein variants are modified compared to the parental virus protein.

In one embodiment, in the modified amino acid sequence amino acid positions in the one or more virus protein variants which are modified compared to the wildtype virus protein and/or the parental virus protein are modified.

In one embodiment, one or more (e.g., all) of the one or more virus protein variants are the virus proteins of one or more virus strains.

In one embodiment, one or more (e.g., all) of the one or more virus strains are natural isolates, or one or more (e.g., all) of the one or more virus strains are mutants of a natural isolate.

In one embodiment, one or more (e.g., all) of the one or more virus strains are virus variant strains that are prevalent or rapidly spreading.

In one embodiment, one or more (e.g., all) of the one or more virus strains are virus variant strains that are variants of concern.

In one embodiment, the parental virus strain and the one or more virus strains are virus variant strains that are prevalent or rapidly spreading.

In one embodiment, the parental virus strain and the one or more virus strains are virus variant strains that are variants of concern.

In one embodiment, the parental virus protein and the one or more virus protein variants are modified compared to the wildtype virus protein.

In one embodiment, the one or more virus protein variants comprise virus protein variants of at least two virus strains.

In one embodiment, the one or more virus protein variants in b) are different from the one or more virus protein variants in c).

In one embodiment, in the modified amino acid sequence amino acid modifications in the parental virus protein compared to the wildtype virus protein do not interfere with amino acid modifications in the modified amino acid positions.

In one embodiment, in the modified amino acid sequence amino acid modifications in the parental virus protein compared to the wildtype virus protein are not in close spatial distance to modified amino acid positions or amino acid modifications in the modified amino acid positions.

In one embodiment, in the modified amino acid sequence modifications in the modified amino acid positions do not result in major structural rearrangements.

In one embodiment, the amino acids in the modified amino acid positions are surface exposed.

In one embodiment, the modified amino acid positions comprise one or more, e.g., at least two amino acid positions. The modifications may correspond to one or more virus protein variants/virus strains and may encompass all modifications present in the one or more virus protein variants/virus strains compared to the parental virus protein/virus strain, or a portion thereof.

In one embodiment, the modified amino acid positions in b) and c) each comprise one or more, e.g., at least two amino acid positions.

In one aspect, the invention relates to a method comprising the steps:

a) providing a nucleic acid comprising a first nucleotide sequence encoding an amino acid sequence comprising at least a fragment of a parental virus protein, wherein amino acid

positions in the at least a fragment of a parental virus protein are modified to comprise amino acids found in the corresponding amino acid positions of one or more SARS-CoV-2 S protein variants; and

b) providing a nucleic acid comprising a second nucleotide sequence encoding an amino acid sequence comprising at least a fragment of a parental virus protein, wherein amino acid positions in the at least a fragment of a parental virus protein are modified to comprise amino acids found in the corresponding amino acid positions of one or more virus protein variants, wherein the amino acid modifications in b) at least partially differ from the amino acid modifications in a).

In one embodiment, the method may comprise one or more further such steps of providing a nucleic acid encoding at least partially different amino acid modifications.

In one embodiment, the nucleic acid is RNA.

Further embodiments are as described herein.

In one aspect, the invention relates to a medical preparation comprising:

a) a nucleic acid comprising a first nucleotide sequence encoding an amino acid sequence comprising at least a fragment of a parental virus protein, wherein amino acid positions in the at least a fragment of a parental virus protein are modified to comprise amino acids found in the corresponding amino acid positions of one or more virus protein variants; and

b) a nucleic acid comprising a second nucleotide sequence encoding an amino acid sequence comprising at least a fragment of a parental virus protein, wherein amino acid positions in the at least a fragment of a parental virus protein are modified to comprise amino acids found in the corresponding amino acid positions of one or more virus protein variants,

wherein the amino acid modifications in b) at least partially differ from the amino acid modifications in a).

In one embodiment, the medical preparation may comprise one or more further nucleic acids encoding at least partially different amino acid modifications.

In one embodiment, the nucleic acid is RNA.

In one embodiment, the RNA is formulated in lipid nanoparticles (LNP).

In one embodiment, the medical preparation is a pharmaceutical composition.

In one embodiment, the medical preparation is a vaccine.

In one embodiment, the medical preparation is a kit.

In one embodiment, the medical preparation further comprises instructions for use of the medical preparation for vaccination against infection with the virus.

In one aspect, the invention relates to the medical preparation for pharmaceutical use.

In one embodiment, the pharmaceutical use comprises vaccination against infection with the virus.

In one aspect, the invention relates to a method of inducing an immune response against the virus in a subject comprising administering to the subject the medical preparation.

In one embodiment, the method is a method for prophylactic treatment against infection with the virus.

In one embodiment, the method is a method for vaccination against infection with the virus.

Further embodiments of the above aspects are as described herein.

In one embodiment, the virus described herein is SARS-CoV-2. In one embodiment, the virus protein described herein is SARS-CoV-2 spike protein (S protein).

In one aspect, the invention relates to a method comprising the steps:

a) identifying amino acid positions in a parental SARS-CoV-2 spike protein (S protein) which are modified compared to the corresponding amino acid positions of one or more SARS-CoV-2 S protein variants; and

b) providing an amino acid sequence comprising at least a fragment of the parental SARS-CoV-2 S protein, wherein amino acid positions in the at least a fragment of a parental SARS-CoV-2 S protein are modified to comprise amino acids found in the corresponding amino acid positions of one or more SARS-CoV-2 S protein variants, or a nucleotide sequence encoding the modified amino acid sequence.

In one embodiment, the method comprises repeating step b) to provide two or more of the modified amino acid sequences, or two or more of the nucleotide sequences encoding two or more of the modified amino acid sequences.

In one embodiment, the two or more modified amino acid sequences are based on the same parental SARS-CoV-2 S protein.

In one embodiment, the amino acid modifications in the two or more modified amino acid sequences are at least partially different.

In one embodiment, providing the nucleotide sequence comprises:

b') substituting codons of a nucleotide sequence encoding an amino acid sequence comprising at least a fragment of a parental SARS-CoV-2 S protein with other codons to obtain a mutated nucleotide sequence that encodes a modified amino acid sequence, wherein amino acid positions in the at least a fragment of a parental SARS-CoV-2 S protein are modified to comprise amino acids found in the corresponding amino acid positions of one or more SARS-CoV-2 S protein variants.

In one embodiment, the method comprises repeating step b') to provide two or more of the mutated nucleotide sequences encoding two or more of the modified amino acid sequences.

In one embodiment, the two or more modified amino acid sequences are based on the same parental SARS-CoV-2 S protein.

In one embodiment, the amino acid modifications in the two or more modified amino acid sequences are at least partially different.

In one aspect, the invention relates to a method comprising the steps:

a) identifying amino acid positions in a parental SARS-CoV-2 spike protein (S protein) which are modified compared to the corresponding amino acid positions of one or more SARS-CoV-2 S protein variants;

b) substituting codons of a first nucleotide sequence encoding an amino acid sequence comprising at least a fragment of a parental SARS-CoV-2 S protein with other codons to obtain a first mutated nucleotide sequence that encodes a modified amino acid sequence, wherein amino acid positions in the at least a fragment of a parental SARS-CoV-2 S protein are modified to comprise amino acids found in the corresponding amino acid positions of one or more SARS-CoV-2 S protein variants; and

c) substituting codons of a second nucleotide sequence encoding an amino acid sequence comprising at least a fragment of a parental SARS-CoV-2 S protein with other codons to obtain a second mutated nucleotide sequence that encodes a modified amino acid sequence, wherein amino acid positions in the at least a fragment of a parental SARS-CoV-2 S protein are

CA 03215771 2023-10-17

modified to comprise amino acids found in the corresponding amino acid positions of one or more SARS-CoV-2 S protein variants,

wherein the amino acid modifications in b) at least partially differ from the amino acid modifications in c).

In one embodiment, the method may comprise one or more further such steps of substituting codons to obtain a mutated nucleotide sequence encoding at least partially different amino acid modifications.

In one embodiment, the at least a fragment of a parental SARS-CoV-2 S protein comprised in the amino acid sequence encoded by the first nucleotide sequence and the at least a fragment of a parental SARS-CoV-2 S protein comprised in the amino acid sequence encoded by the second nucleotide sequence are identical.

In one embodiment, the amino acid sequence encoded by the first nucleotide sequence and the amino acid sequence encoded by the second nucleotide sequence are identical.

In one embodiment, the first nucleotide sequence and the second nucleotide sequence are identical.

In one embodiment, one or more of the modified amino acid positions in the modified amino acid sequence encoded by the first mutated nucleotide sequence differ from the modified amino acid positions in the modified amino acid sequence encoded by the second mutated nucleotide sequence.

In one embodiment, one or more amino acids in modified amino acid positions modified in the modified amino acid sequence encoded by the first mutated nucleotide sequence and in the modified amino acid sequence encoded by the second mutated nucleotide sequence differ from each other.

In one embodiment, the amino acid sequence comprising at least a fragment of a parental SARS-CoV-2 S protein comprises the amino acid sequence of the N-terminal domain (NTD) and/or receptor binding domain (RBD) of a SARS-CoV-2 S protein.

In one embodiment, the amino acid sequence comprising at least a fragment of a parental SARS-CoV-2 S protein comprises the amino acid sequence of a full-length SARS-CoV-2 S protein.

In one embodiment, the method further comprises providing nucleic acid comprising the nucleotide sequence encoding a modified amino acid sequence.

In one embodiment, the method further comprises providing a vaccine comprising nucleic acid comprising the nucleotide sequence encoding a modified amino acid sequence.

In one embodiment, the nucleic acid is RNA.

In one embodiment, the method is a method of generating a SARS-CoV-2 vaccine.

In one embodiment, the vaccine is an RNA vaccine.

In one embodiment, the vaccine has a reduced risk for immune escape.

In one embodiment, one or more of the modified amino acid positions are located within the N-terminal domain (NTD) and/or the receptor binding domain (RBD) of a SARS-CoV-2 S protein.

In one embodiment, the modified amino acid positions are amino acid positions at which the amino acid sequence of the one or more SARS-CoV-2 S protein variants differs from the amino acid sequence of the parental SARS-CoV-2 S protein.

In one embodiment, the modified amino acid positions are amino acid positions at which the amino acid sequence of the one or more SARS-CoV-2 S protein variants differs from the amino acid sequence of wildtype SARS-CoV-2 S protein.

In one embodiment, the modified amino acid positions are potential sites for escape mutants of SARS-CoV-2.

In one embodiment, the escape mutants of SARS-CoV-2 are antibody escape mutants of SARS-CoV-2.

In one embodiment, the escape mutants of SARS-CoV-2 are resistant to neutralization by antibody against SARS-CoV-2 S protein.

In one embodiment, the SARS-CoV-2 S protein of the escape mutants of SARS-CoV-2 shows reduced antibody binding.

In one embodiment, the antibody is used for treating a patient infected with SARS-CoV-2.

In one embodiment, the antibody is generated in a patient that has been treated with a SARS-CoV-2 vaccine.

In one embodiment, the parental SARS-CoV-2 S protein is modified compared to wildtype SARS-CoV-2 S protein.

CA 03215771 2023-10-17

In one embodiment, in the modified amino acid sequence amino acid positions in the parental SARS-CoV-2 S protein which are modified compared to wildtype SARS-CoV-2 S protein are not modified.

In one embodiment, the parental SARS-CoV-2 S protein is the S protein of a parental SARS-CoV-2 strain.

In one embodiment, the parental SARS-CoV-2 strain is a natural isolate, or the parental SARS-CoV-2 strain is a mutant of a natural isolate.

In one embodiment, the parental SARS-CoV-2 strain is a SARS-CoV-2 variant strain that is prevalent or rapidly spreading.

In one embodiment, the parental SARS-CoV-2 strain is a SARS-CoV-2 variant that is a variant of concern.

In one embodiment, the parental SARS-CoV-2 strain is B.1.1.7.

In one embodiment, the one or more SARS-CoV-2 S protein variants are modified compared to wildtype SARS-CoV-2 S protein.

In one embodiment, the one or more SARS-CoV-2 S protein variants are modified compared to the parental SARS-CoV-2 S protein.

In one embodiment, in the modified amino acid sequence amino acid positions in the one or more SARS-CoV-2 S protein variants which are modified compared to wildtype SARS-CoV-2 S protein and/or the parental SARS-CoV-2 S protein are modified.

In one embodiment, one or more (e.g., all) of the one or more SARS-CoV-2 S protein variants are the S proteins of one or more SARS-CoV-2 strains.

In one embodiment, one or more (e.g., all) of the one or more SARS-CoV-2 strains are natural isolates, or one or more (e.g., all) of the one or more SARS-CoV-2 strains are mutants of a natural isolate.

In one embodiment, one or more (e.g., all) of the one or more SARS-CoV-2 strains are SARS-CoV-2 variant strains that are prevalent or rapidly spreading.

In one embodiment, one or more (e.g., all) of the one or more SARS-CoV-2 strains are SARS-CoV-2 variant strains that are variants of concern.

In one embodiment, one or more of the one or more SARS-CoV-2 strains are selected from the group consisting of B.1.351, B.1.1.298, B.1.427/B.1.429, B.1.526, and P1.

CA 03215771 2023-10-17

In one embodiment, the parental SARS-CoV-2 strain and the one or more SARS-CoV-2 strains are SARS-CoV-2 variant strains that are prevalent or rapidly spreading.

In one embodiment, the parental SARS-CoV-2 strain and the one or more SARS-CoV-2 strains are SARS-CoV-2 variant strains that are variants of concern.

In one embodiment, the parental SARS-CoV-2 S protein and the one or more SARS-CoV-2 S protein variants are modified compared to wildtype SARS-CoV-2 S protein.

In one embodiment, the parental SARS-CoV-2 strain is B.1.1.7 and the one or more SARS-CoV-2 strains are selected from the group consisting of B.1.351, B.1.1.298, B.1.427/B.1.429, B.1.526, P1.

In one embodiment, the one or more SARS-CoV-2 S protein variants comprise SARS-CoV-2 S protein variants of at least two SARS-CoV-2 strains.

In one embodiment, the one or more SARS-CoV-2 S protein variants in b) are different from the one or more SARS-CoV-2 S protein variants in c).

In one embodiment, the one or more SARS-CoV-2 S protein variants in b) are the SARS-CoV-2 S protein variants of B.1.427/B.1.429, and B.1.526, and the one or more SARS-CoV-2 S protein variants in c) are the SARS-CoV-2 S protein variants of B.1.351, P.1, and B.1.1.298.

In one embodiment, in the modified amino acid sequence amino acid modifications in the parental SARS-CoV-2 S protein compared to wildtype SARS-CoV-2 S protein do not interfere with amino acid modifications in the modified amino acid positions.

In one embodiment, in the modified amino acid sequence amino acid modifications in the parental SARS-CoV-2 S protein compared to wildtype SARS-CoV-2 S protein are not in close spatial distance to modified amino acid positions or amino acid modifications in the modified amino acid positions.

In one embodiment, in the modified amino acid sequence modifications in the modified amino acid positions do not result in major structural rearrangements.

In one embodiment, the amino acids in the modified amino acid positions are surface exposed.

In one embodiment, the modified amino acid positions comprise one or more, e.g., at least two amino acid positions. The modifications may correspond to one or more SARS-CoV-2 S protein variants/SARS-CoV-2 strains and may encompass all modifications present in the one

or more SARS-CoV-2 S protein variants/SARS-CoV-2 strains compared to the parental SARS-CoV-2 S protein/SARS-CoV-2 strain, or a portion thereof.

In one embodiment, the modified amino acid positions in b) and c) each comprise one or more, e.g., at least two amino acid positions.

In one embodiment, the modified amino acid positions comprise two or more selected from the group consisting of:

18, 20, 26, 80, 138, 144, 190, 215, 246, 253, 417, 439, 452, 453, 477, 484, 501, 570, 701, 716,

140, 345, 346, 352, 378, 406, 420, 440, 441, 444, 445, 446, 450, 455, 460, 475, 478, 485, 486, 487, 489, 490, 493, 494, 499,

142, 145, 146, 147, 150, 152, 154, 156, 157, 158, 164, 247, 248, 249, 250, 251, 252, 254, 255, 258, 365, 369, 370, 374, 376, 384, 405, 408, 415, 421, 443, 447, 448, 456, 472, 473, 476, 496, 498, 500, 504.

In one embodiment, the modifications in the modified amino acid positions comprise two or more selected from the group consisting of:

18F, 20N, 26S, 80Y, 138Y, 144F, 190S, 215A, 246I, 253G, 417N, 439K, 452R, 453F, 477N, 484K, 501Y, 570D, 701V, 716I,

140L, 345A, 346K, 352S, 378N, 406Q, 420, 440K, 441F, 444, 445A, 446V, 450K, 455F, 460I, 475V, 478I, 485V, 486L, 487D, 489, 490S, 493L, 494P, 499H,

142S, 145H, 146Y, 147N, 150R, 152C, 154Q, 156A, 157L, 158G, 164T, 247G, 248H, 249S, 250N, 251S, 252V, 254F, 255F, 258L, 365D, 369C, 370S, 374L, 376I, 384L, 405Y, 408I, 415N, 421, 443A, 447V, 448Y, 456L, 472V, 473F, 476S, 496C, 498H, 500I, 504D.

In one embodiment, the modifications in the modified amino acid positions comprise two or more selected from the group consisting of:

L18F, T20N, P26S, D80Y, D138Y, Y144F, R190S, D215A, R246I, D253G, K417N, N439K, L452R, Y453F, S477N, E484K, N501Y, A570D, A701V, T716I,

F140L, T345A, R346K, A352S, K378N, E406Q, D420, N440K, L441F, K444, V445A, G446V, N450K, L455F, N460I, A475V, T478I, G485V, F486L, N487D, Y489, F490S, Q493L, S494P, P499H,

G142S, Y145H, H146Y, K147N, K150R, W152C, E154Q, E156A, F157L, R158G, N164T, S247G, Y248H, L249S, T250N, P251S, G252V, S254F, S255F, W258L, Y365D, Y369C, N370S, F374L,

CA 03215771 2023-10-17

T376I, P384L, D405Y, R408I, T415N, Y421, S443A, G447V, N448Y, F456L, I472V, Y473F, G476S, G496C, Q498H, T500I, G504D.

In one aspect, the invention relates to a method comprising the steps:

a) providing a nucleic acid comprising a first nucleotide sequence encoding an amino acid sequence comprising at least a fragment of a parental SARS-CoV-2 S protein, wherein amino acid positions in the at least a fragment of a parental SARS-CoV-2 S protein are modified to comprise amino acids found in the corresponding amino acid positions of one or more SARS-CoV-2 S protein variants; and

b) providing a nucleic acid comprising a second nucleotide sequence encoding an amino acid sequence comprising at least a fragment of a parental SARS-CoV-2 S protein, wherein amino acid positions in the at least a fragment of a parental SARS-CoV-2 S protein are modified to comprise amino acids found in the corresponding amino acid positions of one or more SARS-CoV-2 S protein variants,

wherein the amino acid modifications in b) at least partially differ from the amino acid modifications in a).

In one embodiment, the method may comprise one or more further such steps of providing a nucleic acid encoding at least partially different amino acid modifications.

In one embodiment, the nucleic acid is RNA.

Further embodiments are as described herein.

In one aspect, the invention relates to a medical preparation comprising:

a) a nucleic acid comprising a first nucleotide sequence encoding an amino acid sequence comprising at least a fragment of a parental SARS-CoV-2 S protein, wherein amino acid positions in the at least a fragment of a parental SARS-CoV-2 S protein are modified to comprise amino acids found in the corresponding amino acid positions of one or more SARS-CoV-2 S protein variants; and

b) a nucleic acid comprising a second nucleotide sequence encoding an amino acid sequence comprising at least a fragment of a parental SARS-CoV-2 S protein, wherein amino acid positions in the at least a fragment of a parental SARS-CoV-2 S protein are modified to comprise amino acids found in the corresponding amino acid positions of one or more SARS-CoV-2 S protein variants,

wherein the amino acid modifications in b) at least partially differ from the amino acid modifications in a).

In one embodiment, the medical preparation may comprise one or more further nucleic acids encoding at least partially different amino acid modifications.

In one embodiment, the nucleic acid is RNA.

In one embodiment, the RNA is formulated in lipid nanoparticles (LNP).

In one embodiment, the medical preparation is a pharmaceutical composition.

In one embodiment, the medical preparation is a vaccine.

In one embodiment, the medical preparation is a kit.

In one embodiment, the medical preparation further comprises instructions for use of the medical preparation for vaccination against SARS-CoV-2 infection.

In one aspect, the invention relates to the medical preparation for pharmaceutical use.

In one embodiment, the pharmaceutical use comprises vaccination against SARS-CoV-2 infection.

In one aspect, the invention relates to a method of inducing an immune response against SARS-CoV-2 in a subject comprising administering to the subject the medical preparation.

In one embodiment, the method is a method for prophylactic treatment against SARS-CoV-2 infection.

In one embodiment, the method is a method for vaccination against SARS-CoV-2 infection.

Further embodiments of the above aspects are as described herein.

The nucleic acid described herein may be single-stranded RNA and the medical preparation, e.g., vaccine, described herein may comprise as the active principle single-stranded RNA that may be translated into the respective protein upon entering cells of a recipient. In addition to wildtype or codon-optimized sequences encoding the antigen sequence, the RNA may contain one or more structural elements optimized for maximal efficacy of the RNA with respect to stability and translational efficiency (5' cap, 5' UTR, 3' UTR, poly(A)-tail). In one embodiment, the RNA contains all of these elements. In one embodiment, beta-S-ARCA(D1) ($m_2^{7,2'-O}$GppSpG) or $m_2^{7,3'-O}$Gppp($m_1^{2'-O}$)ApG may be utilized as specific capping structure at the 5'-end of the RNA drug substances. As 5'-UTR sequence, the 5'-UTR sequence of the human

CA 03215771 2023-10-17

alpha-globin mRNA, optionally with an optimized 'Kozak sequence' to increase translational efficiency (e.g., SEQ ID NO: 12) may be used. As 3'-UTR sequence, a combination of two sequence elements (FI element) derived from the "amino terminal enhancer of split" (AES) mRNA (called F) and the mitochondrial encoded 12S ribosomal RNA (called I) (e.g., SEQ ID NO: 13) placed between the coding sequence and the poly(A)-tail to assure higher maximum protein levels and prolonged persistence of the mRNA may be used. These were identified by an *ex vivo* selection process for sequences that confer RNA stability and augment total protein expression (see WO 2017/060314, herein incorporated by reference). Alternatively, the 3'-UTR may be two re-iterated 3'-UTRs of the human beta-globin mRNA. Furthermore, a poly(A)-tail measuring 110 nucleotides in length, consisting of a stretch of 30 adenosine residues, followed by a 10 nucleotide linker sequence (of random nucleotides) and another 70 adenosine residues (e.g., SEQ ID NO: 14) may be used. This poly(A)-tail sequence was designed to enhance RNA stability and translational efficiency.

Furthermore, a secretory signal peptide (sec) may be fused to the antigen-encoding regions preferably in a way that the sec is translated as N terminal tag. In one embodiment, sec corresponds to the secreotory signal peptide of the S protein. Sequences coding for short linker peptides predominantly consisting of the amino acids glycine (G) and serine (S), as commonly used for fusion proteins may be used as GS/Linkers.

The RNA described herein may be complexed with proteins and/or lipids, preferably lipids, to generate RNA-particles for administration. If a combination of different RNAs is used, the RNAs may be complexed together or complexed separately with proteins and/or lipids to generate RNA-particles for administration.

In one embodiment, the vaccine antigens comprising polyspecific SARS-CoV-2 S protein amino acid sequences described herein are able to form a multimeric complex, in particular a trimeric complex. In one embodiment, vaccine antigens comprising different polyspecific SARS-CoV-2 S protein amino acid sequences described herein are able to form a multimeric complex, in particular a trimeric complex. Thus, in embodiments of the invention which comprise provision of vaccine antigens comprising different polyspecific SARS-CoV-2 S protein amino acid sequences described herein to a subject, e.g., by administering different nucleic acids

CA 03215771 2023-10-17

encoding the vaccine antigens comprising different polyspecific SARS-CoV-2 S protein amino acid sequences, the different polyspecific SARS-CoV-2 S protein amino acid sequences may be able to form a multimeric complex, in particular a trimeric complex. To this end, the polyspecific SARS-CoV-2 S protein amino acid sequences described herein may comprise a domain allowing the formation of a multimeric complex, in particular a trimeric complex of the polyspecific SARS-CoV-2 S protein amino acid sequences. In one embodiment, the domain allowing the formation of a multimeric complex comprises a trimerization domain, for example, a trimerization domain as described herein, e.g., SARS-CoV-2 S protein trimerization domain. In one embodiment, trimerization is achieved by addition of a trimerization domain, e.g., a T4-fibritin-derived "foldon" trimerization domain (e.g., SEQ ID NO: 10), to the polyspecific SARS-CoV-2 S protein amino acid sequence, in particular if the polyspecific SARS-CoV-2 S protein amino acid sequence corresponds to a portion of a SARS-CoV-2 S protein that does not comprise the SARS-CoV-2 S protein trimerization domain.

In one embodiment, a vaccine antigen comprising a polyspecific virus protein amino acid sequence described herein is encoded by a coding sequence which is codon-optimized and/or the G/C content of which is increased compared to wild type coding sequence, wherein the codon-optimization and/or the increase in the G/C content preferably does not change the sequence of the encoded amino acid sequence.

In one embodiment, a vaccine antigen comprising a polyspecific virus protein amino acid sequence is encoded by RNA. In a particularly preferred embodiment, a vaccine antigen comprising a polyspecific virus protein amino acid sequence is encoded by an isolated messenger ribonucleic acid (mRNA) polynucleotide, wherein the isolated mRNA polynucleotide comprises an open reading frame encoding a polypeptide that comprises the vaccine antigen. In one embodiment, the isolated mRNA polynucleotide is formulated in at least one lipid nanoparticle. For example, in some embodiments, such a lipid nanoparticle may comprise a molar ratio of 20-60% ionizable cationic lipid, 5-25% non-cationic lipid (e.g., neutral lipid), 25-55% sterol or steroid, and 0.5-15% polymer-conjugated lipid (e.g., PEG-modified lipid). In some embodiments, a sterol or steroid included in a lipid nanoparticle may be or comprise cholesterol. In some embodiments, a neutral lipid may be or comprise 1,2-distearoyl-sn-glycero-3-phosphocholine (DSPC). In some embodiments, a polymer-conjugated

CA 03215771 2023-10-17

lipid may be or comprise PEG2000 DMG. In some embodiments, an immunogenic composition may comprise a total lipid content of about 1 mg to 10 mg, or 3 mg to 8 mg, or 4 mg to 6 mg. In some embodiments, such an immunogenic composition may comprise a total lipid content of about 5 mg/mL-15 mg/mL or 7.5 mg/mL-12.5 mg/mL or 9-11 mg/mL. In some embodiments, such an isolated mRNA polynucleotide is provided in an effective amount to induce an immune response in a subject administered at least one dose of the immunogenic composition. In some embodiments, such an isolated mRNA polynucleotide provided in an immunogenic composition is not self-replicating RNA.

In one embodiment, the RNA described herein is a modified RNA, in particular a stabilized mRNA. In one embodiment, the RNA comprises a modified nucleoside in place of at least one uridine. In one embodiment, the RNA comprises a modified nucleoside in place of each uridine. In one embodiment, the modified nucleoside is independently selected from pseudouridine (ψ), N1-methyl-pseudouridine (m1ψ), and 5-methyl-uridine (m5U).

In one embodiment, the RNA comprises a modified nucleoside in place of uridine.

In one embodiment, the modified nucleoside is selected from pseudouridine (ψ), N1-methyl-pseudouridine (m1ψ), and 5-methyl-uridine (m5U).

In one embodiment, the RNA comprises a 5' cap.

In one embodiment, the RNA comprises a 5' UTR comprising the nucleotide sequence of SEQ ID NO: 12, or a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 12.

In one embodiment, the RNA comprises a 3' UTR comprising the nucleotide sequence of SEQ ID NO: 13, or a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 13.

In one embodiment, the RNA comprises a poly-A sequence.

In one embodiment, the poly-A sequence comprises at least 100 nucleotides.

In one embodiment, the poly-A sequence comprises or consists of the nucleotide sequence of SEQ ID NO: 14.

In one embodiment, the RNA is formulated or is to be formulated as a liquid, a solid, or a combination thereof.

In one embodiment, the RNA is formulated or is to be formulated for injection.

CA 03215771 2023-10-17

In one embodiment, the RNA is formulated or is to be formulated for intramuscular administration.

In one embodiment, the RNA is formulated or is to be formulated as particles.

In one embodiment, the particles are lipid nanoparticles (LNP) or lipoplex (LPX) particles.

In one embodiment, the LNP particles comprise ((4-hydroxybutyl)azanediyl)bis(hexane-6,1-diyl)bis(2-hexyldecanoate), 2-[(polyethylene glycol)-2000]-N,N-ditetradecylacetamide, 1,2-Distearoyl-sn-glycero-3-phosphocholine, and cholesterol.

In one embodiment, the RNA lipoplex particles are obtainable by mixing the RNA with liposomes. In one embodiment, the RNA lipoplex particles are obtainable by mixing the RNA with lipids.

In one embodiment, the RNA is formulated or is to be formulated as colloid. In one embodiment, the RNA is formulated or is to be formulated as particles, forming the dispersed phase of a colloid. In one embodiment, 50% or more, 75% or more, or 85% or more of the RNA are present in the dispersed phase. In one embodiment, the RNA is formulated or is to be formulated as particles comprising RNA and lipids. In one embodiment, the particles are formed by exposing RNA, dissolved in an aqueous phase, with lipids, dissolved in an organic phase. In one embodiment, the organic phase comprises ethanol. In one embodiment, the particles are formed by exposing RNA, dissolved in an aqueous phase, with lipids, dispersed in an aqueous phase. In one embodiment, the lipids dispersed in an aqueous phase form liposomes.

In one embodiment, the RNA is mRNA or saRNA.

In one embodiment, the composition or medical preparation described herein is a pharmaceutical composition.

In one embodiment, the composition or medical preparation described herein is a vaccine.

In one embodiment, the pharmaceutical composition further comprises one or more pharmaceutically acceptable carriers, diluents and/or excipients.

In one embodiment, the composition or medical preparation described herein is a kit.

In one embodiment, the RNA and optionally the particle forming components are in separate vials.

In one embodiment, the kit further comprises instructions for use of the composition or medical preparation for inducing an immune response against a virus, e.g., coronavirus, in a subject.

In one aspect, the invention relates to the composition or medical preparation described herein for pharmaceutical use.

In one embodiment, the pharmaceutical use comprises inducing an immune response against a virus, e.g., coronavirus, in a subject.

In one embodiment, the pharmaceutical use comprises a therapeutic or prophylactic treatment of a virus infection, e.g., coronavirus infection.

In one embodiment, the composition or medical preparation described herein is for administration to a human.

In one embodiment, the coronavirus is a betacoronavirus.

In one embodiment, the coronavirus is a sarbecovirus.

In one embodiment, the coronavirus is SARS-CoV-2.

In one aspect, the invention relates to a method of inducing an immune response against coronavirus in a subject comprising administering to the subject a composition described herein, e.g., a vaccine antigen comprising a polyspecific SARS-CoV-2 S protein amino acid sequence or a nucleic acid, e.g., RNA, encoding the vaccine antigen. In one embodiment, the method comprises administration of a plurality of such vaccine antigens or nucleic acids.

In one embodiment, the method is a method for vaccination against coronavirus.

In one embodiment, the method is a method for therapeutic or prophylactic treatment of a coronavirus infection.

In one embodiment, the subject is a human.

In one embodiment, the coronavirus is a betacoronavirus.

In one embodiment, the coronavirus is a sarbecovirus.

In one embodiment, the coronavirus is SARS-CoV-2.

In one embodiment of the method described herein, the composition is a composition described herein.

In one aspect, the invention relates to a composition or medical preparation described herein for use in a method described herein.

In one aspect, the invention relates to a composition, e.g., protein or nucleic acid composition, obtainable by practicing a method described herein

Brief description of the drawings

Figure 1: Schematic overview of the S protein organization of the SARS-CoV-2 S protein.
The sequence within the S1 subunit consists of the signal sequence (SS) and the receptor binding domain (RBD) which is the key subunit within the S protein which is relevant for binding to the human cellular receptor ACE2. The S2 subunit contains the S2 protease cleavage site (S2') followed by a fusion peptide (FP) for membrane fusion, heptad repeats (HR1 and HR2) with a central helix (CH) domain, the transmembrane domain (TM) and a cytoplasmic tail (CT).

Figure 2: Constructs for the development of a SARS-CoV-2 vaccine.
Compositions described herein may be based on different constructs which have been designed based on the full and wildtype S protein comprising the (1) full protein with mutations in close distance to the first heptad repeat (HRP1) that include stabilizing mutations preserving neutralisation sensitive sites, the (2) S1 domain or the (3) RB domain (RBD) only. Furthermore, to stabilize the protein fragments a fibritin domain (F) was fused to the C-terminus. All constructs start with the signal peptide (SP) to ensure Golgi transport to the cell membrane.

Figure 3: General structure of the RNA on the basis of which constructs described herein may be designed.
Schematic illustration of the general structure of the RNA vaccines with 5'-cap, 5'- and 3'-untranslated regions, coding sequences with intrinsic secretory signal peptide as well as GS-linker, and poly(A)-tail. Please note that the individual elements are not drawn exactly true to scale compared to their respective sequence lengths.
UTR = Untranslated region; sec = Secretory signal peptide; RBD = Receptor Binding Domain; GS = Glycine-serine linker.

Figure 4: General structure of the RNA on the basis of which constructs described herein may be designed.

CA 03215771 2023-10-17

Schematic illustration of the general structure of the RNA drug substances with 5'-cap, 5'- and 3'-untranslated regions, coding sequences with intrinsic secretory signal peptide as well as GS-linker, and poly(A)-tail. Please note that the individual elements are not drawn exactly true to scale compared to their respective sequence lengths.
GS = Glycine-serine linker; UTR = Untranslated region; Sec = Secretory signal peptide; RBD = Receptor Binding Domain.

Figure 5: General structure of the RNA on the basis of which constructs described herein may be designed.
Schematic illustration of the general structure of the RNA vaccines with 5'-cap, 5'- and 3'-untranslated regions, coding sequences of the Venezuelan equine encephalitis virus (VEEV) RNA-dependent RNA polymerase replicase and the SARS-CoV-2 antigen with intrinsic secretory signal peptide as well as GS-linker, and poly(A)-tail. Please note that the individual elements are not drawn exactly true to scale compared to their respective sequence lengths.
UTR = Untranslated region; Sec = Secretory signal peptide; RBD = Receptor Binding Domain; GS = Glycine-serine linker.

Figure 6: Exemplary novel bivalent vaccine design based on the B.1.1.7 lineage spike protein
Cryo-EM structure of the SARS-CoV-2 Spike protein with one RBD erect illustrating the surface exposed mutated sites for vaccine sequences (A) S-Seq1 and (B) S-Seq2. The crystal structure of the furin cleaved spike protein of SARS-CoV-2 (PDB ID: 6ZGG) was obtained from the RCSB PDB database and visualized using PyMol v.2.4.1. Sites in yellow indicate amino acid residues that are targeted by lineage B.1.1.7 mutations. Sites in red indicate amino acid residues that are targeted by additional non-B.1.1.7 lineage mutations.

Figure 7: Vaccine candidate expression in HEK293T cells.
Expression of RNA encoded variant SARS-CoV-2 Spike (S) protein in HEK293T cells after transfection of 0.15 µg/mL modRNA using a commercial transfection reagent encoding BNT162b2 and BNT162b2(Alpha+SA) (**A** and **B**), or 0.15 µg/mL LNP-formulated modRNA encoding BNT162b2, BNT162b2(Alpha), and BNT162b2(Alpha;L452R+E484Q) (**C** and **D**),

CA 03215771 2023-10-17

respectively. Surface expression of variant S protein was detected by flow cytometry using a human recombinant ACE-2 protein fused to a mouse Fc-tag and a secondary fluorescent tagged anti-mouse antibody. Percentage of S protein expressing cells (in **A** and **C**) and median fluorescence intensities (MFI) (in **B** and **D**) of the total HEK293T population are depicted per vaccine candidate. Data shown are mean+SD of HEK293T transfections performed in triplicates.

Figure 8: Kinetic of the anti-S1(B.1.1.7) IgG antibody response in sera of Balb/C mice after one immunization with mRNA vaccine candidates encoding different SARS-CoV-2 Alpha variant derived P2 S constructs.

Anti-S1(B.1.1.7) IgG antibody titers after one immunization with 1 µg LNP-formulated modRNA encoding BNT162b2(Alpha), BNT162b2(Alpha+SA) and 0.9% sodium chloride as buffer control at day 7 (**A**), day 14 (**B**), day 21 (**C**) and day 28 (**D**) after immunization. The longitudinal trajectory of the antibody response in serum is shown in **E**. Data are shown as mean ± SEM of all animals measured in duplicates (n=5). Statistical analysis of the different timepoints are given in Table 8.

Figure 9: Kinetic of the anti-RBD(B.1.351) IgG antibody response in sera of Balb/C mice after one immunization with mRNA vaccine candidates encoding different SARS-CoV-2 Alpha variant derived P2 S constructs.

Anti-RBD(B.1.351) IgG antibody titers after one immunization with 1 µg LNP-formulated modRNA encoding BNT162b2(Alpha), BNT162b2(Alpha+SA) and 0.9% sodium chloride as buffer control at day 7 (**A**), day 14 (**B**), day 21 (**C**) and day 28 (**D**) after immunization. The longitudinal trajectory of the antibody response in serum is shown in **E**. Data are shown as mean ± SEM of all animals measured in duplicates (n=5). Statistical analysis of the different timepoints are given in Table 9.

Figure 10: 50% pseudovirus neutralization (pVN_{50}) titers in sera of Balb/C mice at 28 days after one immunization with mRNA vaccine candidates encoding different SARS-CoV-2 Alpha variant derived P2 S constructs.

CA 03215771 2023-10-17

50% pseudovirus neutralizinig antibody (pVN_{50}) titers after one immunization with 1 µg LNP-formulated modRNA encoding BNT162b2(Alpha), BNT162b2(Alpha+SA) and 0.9% sodium chloride as buffer control at 28 days after immunization. The VSV-SARS-CoV-2 pseudoviruses used in the analysis harbor the SARS-CoV-2 S proteins specific for either the ancestral Wuhan strain (Wuhan), the B.1.1.7 (Alpha) variant or the B.1.351 (Beta) variant. Each individual serum (n=5 per group) was tested in duplicate and geometric mean pVN_{50} titers were plotted. For values below the limit of detection (LOD; 12), LOD/2 values are plotted. Group geometric mean titers (horizontal lines and values above bars) with 95% confidence intervals are shown.

Figure 11: 50% pseudovirus neutralization (pVN_{50}) titers in sera of Balb/C mice at 28 days after one immunization with BNT16b2(Alpha) or BNT162b2(Alpha;L452R+E484Q).

50% pseudovirus neutralizinig antibody (pVN_{50}) titers after one immunization with 1 µg LNP-formulated modRNA encoding BNT162b2(Alpha), BNT162b2(Alpha;L452R+E484Q) or 0.9% sodium chloride as buffer control at 28 days after immunization. The VSV-SARS-CoV-2 pseudoviruses used in the analysis harbor the SARS-CoV-2 S proteins specific for either the B.1.1.7 (Alpha) variant, the B.1.617.1 (Kappa) variant or the B.1.617.2 (Delta) variant. Each individual serum (n=5 per group) was tested in duplicate and geometric mean pVN50 titers were plotted. For values below the limit of detection (LOD; 12), LOD/2 values are plotted. Group geometric mean titers (horizontal lines and values above bars) with 95% confidence intervals are shown.

Figure 12: IgG response against different recombinant S proteins at 28 days in sera of Balb/c mice after one immunization with BNT16b2(Alpha) or BNT162b2(Alpha;L452R+E484Q) measured by multiplex analysis.

ECL signal after one immunization with 1 µg LNP-formulated modRNA encoding BNT162b2(Alpha), BNT162b2(Alpha;L452R+E484Q), or 0.9% sodium chloride as buffer control at 28 days after immunization. ECL signal was analyzed for linear range (indicated by dotted lines in the different graphs) and response against SARS-CoV-2 S (B.1.1.7; 1:5400 serum dilution) (**A**), SARS-CoV-2 S (B.1.1.529; BA.1; 1:1800 serum dilution) (**B**), SARS-CoV-2 S (BA.1+L452R; 1:1800 serum dilution) (**C**) are given. Data shown represent serum from

CA 03215771 2023-10-17

individual animals measured in duplicates, horizontal lines indicate the group mean±SEM (n=5 per group; buffer group only n=3). Statistical analysis of the different timepoints are given in Table 12.

Detailed description

Although the present disclosure is described in detail below, it is to be understood that this disclosure is not limited to the particular methodologies, protocols and reagents described herein as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present disclosure which will be limited only by the appended claims. Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art.

Preferably, the terms used herein are defined as described in "A multilingual glossary of biotechnological terms: (IUPAC Recommendations)", H.G.W. Leuenberger, B. Nagel, and H. Kölbl, Eds., Helvetica Chimica Acta, CH-4010 Basel, Switzerland, (1995).

The practice of the present disclosure will employ, unless otherwise indicated, conventional methods of chemistry, biochemistry, cell biology, immunology, and recombinant DNA techniques which are explained in the literature in the field (cf., e.g., Molecular Cloning: A Laboratory Manual, 2nd Edition, J. Sambrook et al. eds., Cold Spring Harbor Laboratory Press, Cold Spring Harbor 1989).

In the following, the elements of the present disclosure will be described. These elements are listed with specific embodiments, however, it should be understood that they may be combined in any manner and in any number to create additional embodiments. The variously described examples and embodiments should not be construed to limit the present disclosure to only the explicitly described embodiments. This description should be understood to disclose and encompass embodiments which combine the explicitly described embodiments with any number of the disclosed elements. Furthermore, any permutations and combinations of all described elements should be considered disclosed by this description unless the context indicates otherwise.

Several documents are cited throughout the text of this specification. Each of the documents cited herein (including all patents, patent applications, scientific publications, manufacturer's specifications, instructions, etc.), whether supra or infra, are hereby incorporated by reference in their entirety. Nothing herein is to be construed as an admission that the present disclosure was not entitled to antedate such disclosure.

CA 03215771 2023-10-17

Definitions

In the following, definitions will be provided which apply to all aspects of the present disclosure. The following terms have the following meanings unless otherwise indicated. Any undefined terms have their art recognized meanings.

The term "about" means approximately or nearly, and in the context of a numerical value or range set forth herein in one embodiment means ± 20%, ± 10%, ± 5%, or ± 3% of the numerical value or range recited or claimed.

The terms "a" and "an" and "the" and similar reference used in the context of describing the disclosure (especially in the context of the claims) are to be construed to cover both the singular and the plural, unless otherwise indicated herein or clearly contradicted by context. Recitation of ranges of values herein is merely intended to serve as a shorthand method of referring individually to each separate value falling within the range. Unless otherwise indicated herein, each individual value is incorporated into the specification as if it was individually recited herein. All methods described herein can be performed in any suitable order unless otherwise indicated herein or otherwise clearly contradicted by context. The use of any and all examples, or exemplary language (e.g., "such as"), provided herein is intended merely to better illustrate the disclosure and does not pose a limitation on the scope of the claims. No language in the specification should be construed as indicating any non-claimed element essential to the practice of the disclosure.

Unless expressly specified otherwise, the term "comprising" is used in the context of the present document to indicate that further members may optionally be present in addition to the members of the list introduced by "comprising". It is, however, contemplated as a specific embodiment of the present disclosure that the term "comprising" encompasses the possibility of no further members being present, i.e., for the purpose of this embodiment "comprising" is to be understood as having the meaning of "consisting of" or "consisting essentially of".

The term "one or more", or "at least one" as used herein refers to one or more than one, including at least two, at least three, at least four, at least five, at least six, at least seven, at least eight, at least nine, at least ten, or even more.

The term "two or more", "at least two", "a plurality", or "poly" refer to more than one including at least two, at least three, at least four, at least five, at least six, at least seven, at least eight, at least nine, at least ten, or even more. Terms such as "reduce", "decrease", "inhibit" or "impair" as used herein relate to an overall reduction or the ability to cause an overall reduction, preferably of at least 5%, at least 10%, at least 20%, at least 50%, at least 75% or even more, in the level. These terms include a complete or essentially complete inhibition, i.e., a reduction to zero or essentially to zero.

The term "polyspecific" when referring to a virus protein amino acid sequence, e.g., SARS-CoV-2 S protein amino acid sequence" refers to an amino acid sequence based on said virus protein wherein multiple epitopes which are normally present in different variants of the virus protein, e.g., different naturally occurring strains, are combined to be present in a single amino acid sequence.

Terms such as "increase", "enhance" or "exceed" preferably relate to an increase or enhancement by at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 80%, at least 100%, at least 200%, at least 500%, or even more.

According to the disclosure, the term "peptide" comprises oligo- and polypeptides and refers to substances which comprise about two or more, about 3 or more, about 4 or more, about 6 or more, about 8 or more, about 10 or more, about 13 or more, about 16 or more, about 20 or more, and up to about 50, about 100 or about 150, consecutive amino acids linked to one another via peptide bonds. The term "protein" or "polypeptide" refers to large peptides, in particular peptides having at least about 150 amino acids, but the terms "peptide", "protein" and "polypeptide" are used herein usually as synonyms.

A "therapeutic protein" has a positive or advantageous effect on a condition or disease state of a subject when provided to the subject in a therapeutically effective amount. In one embodiment, a therapeutic protein has curative or palliative properties and may be administered to ameliorate, relieve, alleviate, reverse, delay onset of or lessen the severity of one or more symptoms of a disease or disorder. A therapeutic protein may have prophylactic properties and may be used to delay the onset of a disease or to lessen the severity of such disease or pathological condition. The term "therapeutic protein" includes entire proteins or peptides, and can also refer to therapeutically active fragments thereof. It can also include

therapeutically active variants of a protein. Examples of therapeutically active proteins include, but are not limited to, antigens for vaccination and immunostimulants such as cytokines.

"Fragment" or "portion", with reference to an amino acid sequence (peptide or protein), relates to a part of an amino acid sequence, i.e. a sequence which represents the amino acid sequence shortened at the N-terminus and/or C-terminus. A fragment shortened at the C-terminus (N-terminal fragment) is obtainable e.g. by translation of a truncated open reading frame that lacks the 3'-end of the open reading frame. A fragment shortened at the N-terminus (C-terminal fragment) is obtainable e.g. by translation of a truncated open reading frame that lacks the 5'-end of the open reading frame, as long as the truncated open reading frame comprises a start codon that serves to initiate translation. A fragment of an amino acid sequence comprises e.g. at least 50 %, at least 60 %, at least 70 %, at least 80%, at least 90% of the amino acid residues from an amino acid sequence. A fragment of an amino acid sequence preferably comprises at least 6, in particular at least 8, at least 12, at least 15, at least 20, at least 30, at least 50, or at least 100 consecutive amino acids from an amino acid sequence.

By "variant" herein is meant an amino acid sequence that differs from a parent amino acid sequence by virtue of at least one amino acid modification. The parent amino acid sequence may be a naturally occurring or wild type (WT) amino acid sequence, or may be a modified version of a wild type amino acid sequence. Preferably, the variant amino acid sequence has at least one amino acid modification compared to the parent amino acid sequence, e.g., from 1 to about 20 amino acid modifications, and preferably from 1 to about 10 or from 1 to about 5 amino acid modifications compared to the parent.

By "wild type" or "WT" or "native" herein is meant an amino acid sequence that is found in nature, including allelic variations. A wild type amino acid sequence, peptide or protein has an amino acid sequence that has not been intentionally modified.

For the purposes of the present disclosure, "variants" of an amino acid sequence (peptide, protein or polypeptide) comprise amino acid insertion variants, amino acid addition variants, amino acid deletion variants and/or amino acid substitution variants. The term "variant" includes all mutants, splice variants, posttranslationally modified variants, conformations,

CA 03215771 2023-10-17

isoforms, allelic variants, species variants, and species homologs, in particular those which are naturally occurring. The term "variant" includes, in particular, fragments of an amino acid sequence.

Amino acid insertion variants comprise insertions of single or two or more amino acids in a particular amino acid sequence. In the case of amino acid sequence variants having an insertion, one or more amino acid residues are inserted into a particular site in an amino acid sequence, although random insertion with appropriate screening of the resulting product is also possible. Amino acid addition variants comprise amino- and/or carboxy-terminal fusions of one or more amino acids, such as 1, 2, 3, 5, 10, 20, 30, 50, or more amino acids. Amino acid deletion variants are characterized by the removal of one or more amino acids from the sequence, such as by removal of 1, 2, 3, 5, 10, 20, 30, 50, or more amino acids. The deletions may be in any position of the protein. Amino acid deletion variants that comprise the deletion at the N-terminal and/or C-terminal end of the protein are also called N-terminal and/or C-terminal truncation variants. Amino acid substitution variants are characterized by at least one residue in the sequence being removed and another residue being inserted in its place. Preference is given to the modifications being in positions in the amino acid sequence which are not conserved between homologous proteins or peptides and/or to replacing amino acids with other ones having similar properties. Preferably, amino acid changes in peptide and protein variants are conservative amino acid changes, i.e., substitutions of similarly charged or uncharged amino acids. A conservative amino acid change involves substitution of one of a family of amino acids which are related in their side chains. Naturally occurring amino acids are generally divided into four families: acidic (aspartate, glutamate), basic (lysine, arginine, histidine), non-polar (alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), and uncharged polar (glycine, asparagine, glutamine, cysteine, serine, threonine, tyrosine) amino acids. Phenylalanine, tryptophan, and tyrosine are sometimes classified jointly as aromatic amino acids. In one embodiment, conservative amino acid substitutions include substitutions within the following groups:

glycine, alanine;

valine, isoleucine, leucine;

aspartic acid, glutamic acid;

CA 03215771 2023-10-17

asparagine, glutamine;

serine, threonine;

lysine, arginine; and

phenylalanine, tyrosine.

Preferably the degree of similarity, preferably identity between a given amino acid sequence and an amino acid sequence which is a variant of said given amino acid sequence will be at least about 60%, 70%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99%. The degree of similarity or identity is given preferably for an amino acid region which is at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90% or about 100% of the entire length of the reference amino acid sequence. For example, if the reference amino acid sequence consists of 200 amino acids, the degree of similarity or identity is given preferably for at least about 20, at least about 40, at least about 60, at least about 80, at least about 100, at least about 120, at least about 140, at least about 160, at least about 180, or about 200 amino acids, in some embodiments continuous amino acids. In some embodiments, the degree of similarity or identity is given for the entire length of the reference amino acid sequence. The alignment for determining sequence similarity, preferably sequence identity can be done with art known tools, preferably using the best sequence alignment, for example, using Align, using standard settings, preferably EMBOSS::needle, Matrix: Blosum62, Gap Open 10.0, Gap Extend 0.5.

"Sequence similarity" indicates the percentage of amino acids that either are identical or that represent conservative amino acid substitutions. "Sequence identity" between two amino acid sequences indicates the percentage of amino acids that are identical between the sequences. "Sequence identity" between two nucleic acid sequences indicates the percentage of nucleotides that are identical between the sequences.

The terms "% identical", "% identity" or similar terms are intended to refer, in particular, to the percentage of nucleotides or amino acids which are identical in an optimal alignment between the sequences to be compared. Said percentage is purely statistical, and the differences between the two sequences may be but are not necessarily randomly distributed over the entire length of the sequences to be compared. Comparisons of two sequences are

usually carried out by comparing the sequences, after optimal alignment, with respect to a segment or "window of comparison", in order to identify local regions of corresponding sequences. The optimal alignment for a comparison may be carried out manually or with the aid of the local homology algorithm by Smith and Waterman, 1981, Ads App. Math. 2, 482, with the aid of the local homology algorithm by Neddleman and Wunsch, 1970, J. Mol. Biol. 48, 443, with the aid of the similarity search algorithm by Pearson and Lipman, 1988, Proc. Natl Acad. Sci. USA 88, 2444, or with the aid of computer programs using said algorithms (GAP, BESTFIT, FASTA, BLAST P, BLAST N and TFASTA in Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Drive, Madison, Wis.). In some embodiments, percent identity of two sequences is determined using the BLASTN or BLASTP algorithm, as available on the United States National Center for Biotechnology Information (NCBI) website (e.g., at blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE_TYPE=BlastSearch&BLAST_SPEC=blast2seq&LINK_LOC=align2seq). In some embodiments, the algorithm parameters used for BLASTN algorithm on the NCBI website include: (i) Expect Threshold set to 10; (ii) Word Size set to 28; (iii) Max matches in a query range set to 0; (iv) Match/Mismatch Scores set to 1, -2; (v) Gap Costs set to Linear; and (vi) the filter for low complexity regions being used. In some embodiments, the algorithm parameters used for BLASTP algorithm on the NCBI website include: (i) Expect Threshold set to 10; (ii) Word Size set to 3; (iii) Max matches in a query range set to 0; (iv) Matrix set to BLOSUM62; (v) Gap Costs set to Existence: 11 Extension: 1; and (vi) conditional compositional score matrix adjustment.

Percentage identity is obtained by determining the number of identical positions at which the sequences to be compared correspond, dividing this number by the number of positions compared (e.g., the number of positions in the reference sequence) and multiplying this result by 100.

In some embodiments, the degree of similarity or identity is given for a region which is at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90% or about 100% of the entire length of the reference sequence. For example, if the reference nucleic acid sequence consists of 200 nucleotides, the degree of identity is given for at least about 100, at least about 120, at least about 140, at least about 160, at least about 180, or

CA 03215771 2023-10-17

about 200 nucleotides, in some embodiments continuous nucleotides. In some embodiments, the degree of similarity or identity is given for the entire length of the reference sequence. Homologous amino acid sequences exhibit according to the disclosure at least 40%, in particular at least 50%, at least 60%, at least 70%, at least 80%, at least 90% and preferably at least 95%, at least 98 or at least 99% identity of the amino acid residues.

The amino acid sequence variants described herein may readily be prepared by the skilled person, for example, by recombinant DNA manipulation. The manipulation of DNA sequences for preparing peptides or proteins having substitutions, additions, insertions or deletions, is described in detail in Sambrook et al. (1989), for example. Furthermore, the peptides and amino acid variants described herein may be readily prepared with the aid of known peptide synthesis techniques such as, for example, by solid phase synthesis and similar methods.

In one embodiment, a fragment or variant of an amino acid sequence (peptide or protein) is preferably a "functional fragment" or "functional variant". The term "functional fragment" or "functional variant" of an amino acid sequence relates to any fragment or variant exhibiting one or more functional properties identical or similar to those of the amino acid sequence from which it is derived, i.e., it is functionally equivalent. With respect to antigens or antigenic sequences, one particular function is one or more immunogenic activities displayed by the amino acid sequence from which the fragment or variant is derived. The term "functional fragment" or "functional variant", as used herein, in particular refers to a variant molecule or sequence that comprises an amino acid sequence that is altered by one or more amino acids compared to the amino acid sequence of the parent molecule or sequence and that is still capable of fulfilling one or more of the functions of the parent molecule or sequence, e.g., inducing an immune response. In one embodiment, the modifications in the amino acid sequence of the parent molecule or sequence do not significantly affect or alter the characteristics of the molecule or sequence. In different embodiments, the function of the functional fragment or functional variant may be reduced but still significantly present, e.g., immunogenicity of the functional variant may be at least 50%, at least 60%, at least 70%, at least 80%, or at least 90% of the parent molecule or sequence. However, in other embodiments, immunogenicity of the functional fragment or functional variant may be enhanced compared to the parent molecule or sequence.

An amino acid sequence (peptide, protein or polypeptide) "derived from" a designated amino acid sequence (peptide, protein or polypeptide) refers to the origin of the first amino acid sequence. Preferably, the amino acid sequence which is derived from a particular amino acid sequence has an amino acid sequence that is identical, essentially identical or homologous to that particular sequence or a fragment thereof. Amino acid sequences derived from a particular amino acid sequence may be variants of that particular sequence or a fragment thereof. For example, it will be understood by one of ordinary skill in the art that the antigens suitable for use herein may be altered such that they vary in sequence from the naturally occurring or native sequences from which they were derived, while retaining the desirable activity of the native sequences.

As used herein, an "instructional material" or "instructions" includes a publication, a recording, a diagram, or any other medium of expression which can be used to communicate the usefulness of the compositions and methods of the invention. The instructional material of the kit of the invention may, for example, be affixed to a container which contains the compositions of the invention or be shipped together with a container which contains the compositions. Alternatively, the instructional material may be shipped separately from the container with the intention that the instructional material and the compositions be used cooperatively by the recipient.

"Isolated" means altered or removed from the natural state. For example, a nucleic acid or a peptide naturally present in a living animal is not "isolated", but the same nucleic acid or peptide partially or completely separated from the coexisting materials of its natural state is "isolated". An isolated nucleic acid or protein can exist in substantially purified form, or can exist in a non-native environment such as, for example, a host cell.

The term "recombinant" in the context of the present invention means "made through genetic engineering". Preferably, a "recombinant object" such as a recombinant nucleic acid in the context of the present invention is not occurring naturally.

The term "naturally occurring" as used herein refers to the fact that an object can be found in nature. For example, a peptide or nucleic acid that is present in an organism (including viruses) and can be isolated from a source in nature and which has not been intentionally modified by man in the laboratory is naturally occurring.

"Physiological pH" as used herein refers to a pH of about 7.5.

The term "genetic modification" or simply "modification" includes the transfection of cells with nucleic acid. The term "transfection" relates to the introduction of nucleic acids, in particular RNA, into a cell. For purposes of the present invention, the term "transfection" also includes the introduction of a nucleic acid into a cell or the uptake of a nucleic acid by such cell, wherein the cell may be present in a subject, e.g., a patient. Thus, according to the present invention, a cell for transfection of a nucleic acid described herein can be present *in vitro* or *in vivo*, e.g. the cell can form part of an organ, a tissue and/or an organism of a patient. According to the invention, transfection can be transient or stable. For some applications of transfection, it is sufficient if the transfected genetic material is only transiently expressed. RNA can be transfected into cells to transiently express its coded protein. Since the nucleic acid introduced in the transfection process is usually not integrated into the nuclear genome, the foreign nucleic acid will be diluted through mitosis or degraded. Cells allowing episomal amplification of nucleic acids greatly reduce the rate of dilution. If it is desired that the transfected nucleic acid actually remains in the genome of the cell and its daughter cells, a stable transfection must occur. Such stable transfection can be achieved by using virus-based systems or transposon-based systems for transfection. Generally, nucleic acid encoding antigen is transiently transfected into cells. RNA can be transfected into cells to transiently express its coded protein.

Coronavirus

Coronaviruses are enveloped, positive-sense, single-stranded RNA ((+) ssRNA) viruses. They have the largest genomes (26–32 kb) among known RNA viruses and are phylogenetically divided into four genera (α, β, γ, and δ), with betacoronaviruses further subdivided into four lineages (A, B, C, and D). Coronaviruses infect a wide range of avian and mammalian species, including humans. Some human coronaviruses generally cause mild respiratory diseases, although severity can be greater in infants, the elderly, and the immunocompromised. Middle East respiratory syndrome coronavirus (MERS-CoV) and severe acute respiratory syndrome coronavirus (SARS-CoV), belonging to betacoronavirus lineages C and B, respectively, are highly pathogenic. Both viruses emerged into the human population from animal reservoirs

CA 03215771 2023-10-17

within the last 15 years and caused outbreaks with high case-fatality rates. The outbreak of severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) that causes atypical pneumonia (coronavirus disease 2019; COVID-19) has raged in China since mid-December 2019, and has developed to be a public health emergency of international concern. SARS-CoV-2 (MN908947.3) belongs to betacoronavirus lineage B. It has at least 70% sequence similarity to SARS-CoV.

In general, coronaviruses have four structural proteins, namely, envelope (E), membrane (M), nucleocapsid (N), and spike (S). The E and M proteins have important functions in the viral assembly, and the N protein is necessary for viral RNA synthesis. The critical glycoprotein S is responsible for virus binding and entry into target cells. The S protein is synthesized as a single-chain inactive precursor that is cleaved by furin-like host proteases in the producing cell into two noncovalently associated subunits, S1 and S2. The S1 subunit contains the receptor-binding domain (RBD), which recognizes the host-cell receptor. The S2 subunit contains the fusion peptide, two heptad repeats, and a transmembrane domain, all of which are required to mediate fusion of the viral and host-cell membranes by undergoing a large conformational rearrangement. The S1 and S2 subunits trimerize to form a large prefusion spike.

The S precursor protein of SARS-CoV-2 can be proteolytically cleaved into S1 (685 aa) and S2 (588 aa) subunits. The S1 subunit comprises the receptor-binding domain (RBD), which mediates virus entry into sensitive cells through the host angiotensin-converting enzyme 2 (ACE2) receptor.

Antigen

Described herein are amino acid sequences comprising sequences which are derived from a virus protein, e.g., a virus surface protein, e.g., SARS-CoV-2 S protein. The virus protein-derived sequences correspond to at least a portion of a virus protein containing amino acid modifications present in other variants of the virus protein, e.g., virus protein variants of other strains. The modification generates (additional) epitopes that are specific for such other virus protein variants. Accordingly, the modified virus protein sequences described herein are polyspecific virus protein amino acid sequences since they contain (modified) amino acid residues that are characteristic for a plurality of virus protein variants and/or virus strains.

CA 03215771 2023-10-17

Also described herein are nucleic acids such as RNA encoding the amino acid sequences comprising virus protein-derived sequences. The amino acid sequences and nucleic acids described herein are useful for inducing an immune response against virus protein, in particular virus protein of different virus strains, in a subject. An amino acid sequence comprising a polyspecific virus protein amino acid sequence described herein (i.e., an antigenic peptide or protein) is also designated herein as "vaccine antigen", "peptide and protein antigen", "antigen molecule" or simply "antigen". The polyspecific virus protein amino acid sequence is also designated herein as "antigenic peptide or protein" or "antigenic sequence".

SARS-CoV-2 coronavirus full length spike (S) protein consist of 1273 amino acids and has the amino acid sequence according to SEQ ID NO: 1:

MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHVSGTNG
TKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLGVYYHKNNK
SWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEP
LVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETK
CTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFS
TFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYN
YLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVC
GPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITP
GTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQ
TQTNSPRRARSVASQSIIAYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECS
NLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNK
VTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAM
QMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISS
VLNDILSRLDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLM
SFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFVSGN
CDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDL
QELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCLKGCCSCGSCCKFDEDDSEPVLKGVKLHYT
(SEQ ID NO: 1)

For purposes of the present disclosure, the above sequence is considered the wildtype SARS-CoV-2 S protein amino acid sequence. Position numberings in SARS-CoV-2 S protein given

CA 03215771 2023-10-17

herein are in relation to the amino acid sequence according to SEQ ID NO: 1 and corresponding positions in SARS-CoV-2 S protein variants.

Full length spike (S) protein according to SEQ ID NO: 1 may be modified in such a way that the prototypical prefusion conformation is stabilized. Stabilization of the prefusion conformation may be obtained by introducing two consecutive proline substitutions at AS residues 986 and 987 in the full length spike protein. Specifically, spike (S) protein stabilized protein variants are obtained in a way that the amino acid residue at position 986 is exchanged to proline and the amino acid residue at position 987 is also exchanged to proline. In one embodiment, a SARS-CoV-2 S protein variant wherein the prototypical prefusion conformation is stabilized comprises the amino acid sequence shown in SEQ ID NO: 7:

MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHVSGTNG
TKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLGVYYHKNNK
SWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEP
LVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETK
CTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFS
TFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYN
YLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVC
GPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITP
GTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQ
TQTNSPRRARSVASQSIIAYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECS
NLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNK
VTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAM
QMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISS
VLNDILSRLDPPEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLM
SFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFVSGN
CDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDL
QELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCLKGCCSCGSCCKFDEDDSEPVLKGVKLHYT
(SEQ ID NO: 7)

In some embodiments, compositions and/or methods described herein are characterized in that sera of vaccinated subjects display neutralizing activity across a panel (e.g., at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 15,

CA 03215771 2023-10-17

or more) of SARS-CoV-2 spike variants. In some embodiments, such SARS-CoV-2 spike variants include mutations in the N-terminal domain (NTD) and/or the receptor binding domain (RBD). In some embodiments, compositions described herein are characterized in that an amino acid sequence comprising a polyspecific SARS-CoV-2 S protein amino acid sequence comprises one or more amino acid modifications (and thus, epitopes) which are characteristic for a panel (e.g., at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 15, or more) of SARS-CoV-2 spike variants. In some embodiments, such modifications which are characteristic for SARS-CoV-2 spike variants include mutations in the N-terminal domain (NTD) and/or the receptor binding domain (RBD).

Modifications in the N-terminal domain (NTD) include modifications at positions selected from the group consisting of:

18, 20, 26, 80, 138, 144, 190, 215, 246, 253,

140,

142, 145, 146, 147, 150, 152, 154, 156, 157, 158, 164, 247, 248, 249, 250, 251, 252, 254, 255, 258.

Specific modifications in the N-terminal domain (NTD) include modifications selected from the group consisting of:

L18F, T20N, P26S, D80Y, D138Y, Y144F, R190S,D215A, R246I, D253G,

F140L,

G142S, Y145H, H146Y, K147N, K150R, W152C, E154Q, E156A, F157L, R158G, N164T, S247G, Y248H, L249S, T250N, P251S, G252V, S254F, S255F, W258L

Modifications in the receptor binding domain (RBD) include modifications at positions selected from the group consisting of:

417, 439, 452, 453, 477, 484, 501,

345, 346, 352, 378, 406, 420, 440, 441, 444, 445, 446, 450, 455, 460, 475, 478, 485, 486, 487, 489, 490, 493, 494, 499,

365, 369, 370, 374, 376, 384, 405, 408, 415, 421, 443, 447, 448, 456, 472, 473, 476, 496, 498, 500, 504.

Specific modifications in the receptor binding domain (RBD) include modifications selected from the group consisting of:

CA 03215771 2023-10-17

K417N, N439K, L452R, Y453F, S477N, E484K, N501Y,
T345A, R346K, A352S, K378N, E406Q, D420, N440K, L441F, K444, V445A, G446V, N450K, L455F, N460I, A475V, T478I, G485V, F486L, N487D, Y489, F490S, Q493L, S494P, P499H,
Y365D, Y369C, N370S, F374L, T376I, P384L, D405Y, R408I, T415N, Y421, S443A, G447V, N448Y, F456L, I472V, Y473F, G476S, G496C, Q498H, T500I, G504D

In some embodiments, compositions described herein are characterized in that an amino acid sequence comprising a polyspecific SARS-CoV-2 S protein amino acid sequence is based on a parental SARS-CoV-2 S protein amino acid sequence (e.g., of a parental SARS-CoV-2 strain) which comprises one or more amino acid modifications compared to wildtype SARS-CoV-2 S protein amino acid sequence, e.g., as compared to SEQ ID NO: 1, and includes one or more further amino acid modifications present in one or more other SARS-CoV-2 S protein amino acid sequences (e.g., of one or more other SARS-CoV-2 strains) compared to wildtype SARS-CoV-2 S protein amino acid sequence, e.g., as compared to SEQ ID NO: 1. Thus, in some embodiments, compositions described herein are characterized in that an amino acid sequence comprising a polyspecific SARS-CoV-2 S protein amino acid sequence may include one or more amino acid modifications present in a parental SARS-CoV-2 S protein amino acid sequence compared to wildtype SARS-CoV-2 S protein amino acid sequence, e.g., as compared to SEQ ID NO: 1, and one or more further amino acid modifications present in one or more other SARS-CoV-2 S protein amino acid sequences compared to wildtype SARS-CoV-2 S protein amino acid sequence, e.g., as compared to SEQ ID NO: 1.

For example, in some embodiments, compositions described herein may be characterized in that an amino acid sequence comprising a polyspecific SARS-CoV-2 S protein amino acid sequence may be based on the SARS-CoV-2 S protein amino acid sequence of strain B.1.1.7 as the parental SARS-CoV-2 S protein amino acid sequence and thus, may include the amino acid modifications present in the parental SARS-CoV-2 S protein amino acid sequence compared to wildtype SARS-CoV-2 S protein amino acid sequence, e.g., as compared to SEQ ID NO: 1, i.e., H69/V70 deletion, Y144 deletion, N501Y, A570D, D614G, P681H, T716I, S982A, and D1118H. In addition, the amino acid sequence comprising a polyspecific SARS-CoV-2 S protein amino acid sequence may include one or more of the amino acid modifications present in one

CA 03215771 2023-10-17

or more other SARS-CoV-2 S protein amino acid sequences, e.g., of one or more strains other than the parental strain, compared to wildtype SARS-CoV-2 S protein amino acid sequence, e.g., as compared to SEQ ID NO: 1. For example, such one or more strains other than the parental strain may be selected from the group consisting of B.1.351, B.1.1.298, B.1.427/B.1.429, B.1.526 and P.1.

In one embodiment, the amino acid sequence comprising a polyspecific SARS-CoV-2 S protein amino acid sequence may include (in addition to the amino acid modifications present in the parental SARS-CoV-2 S protein amino acid sequence of e.g. strain B.1.1.7) one or more of the amino acid modifications present in B.1.526 (NY) and one or more of the amino acid modifications present in B.1.427/B.1.429 (CAL). The one or more of the amino acid modifications present in B.1.526 (NY) may be D253G and A701V. The one or more of the amino acid modifications present in B.1.427/B.1.429 (CAL) may be L452R. Further modifications may include L18F (as present, e.g., in B.1.351 and P.1), N439K and S477N.

In one embodiment, the amino acid sequence comprising a polyspecific SARS-CoV-2 S protein amino acid sequence may include (in addition to the amino acid modifications present in the parental SARS-CoV-2 S protein amino acid sequence of e.g. strain B.1.1.7) one or more of the amino acid modifications present in B.1.351 (SA). The one or more of the amino acid modifications present in B.1.351 (SA) may be D80A, D215G, R246I, K417N and E484K.

Both of the above different amino acid sequences comprising a polyspecific SARS-CoV-2 S protein amino acid sequence (or coding nucleic acids) may be used in combination to induce neutralizing activity across the SARS-CoV-2 spike variants of at least strains B.1.351, B.1.1.298, B.1.427/B.1.429, B.1.526, and P.1.

Those skilled in the art are aware of various spike variants, and/or resources that document them. For example, the following strains, their SARS-CoV-2 S protein amino acid sequences and, in particular, modifications thereof compared to wildtype SARS-CoV-2 S protein amino acid sequence, e.g., as compared to SEQ ID NO: 1, are useful herein.

B.1.1.7 ("Variant of Concern 202012/01" (VOC-202012/01)

B.1.1.7 is a variant of SARS-CoV-2 which was first detected in October 2020 during the COVID-19 pandemic in the United Kingdom from a sample taken the previous month, and it quickly

CA 03215771 2023-10-17

began to spread by mid-December. It is correlated with a significant increase in the rate of COVID-19 infection in United Kingdom; this increase is thought to be at least partly because of change N501Y inside the spike glycoprotein's receptor-binding domain, which is needed for binding to ACE2 in human cells. The B.1.1.7 variant is defined by 23 mutations: 13 non-synonymous mutations, 4 deletions, and 6 synonymous mutations (i.e., there are 17 mutations that change proteins and six that do not). The spike protein changes in B.1.1.7 include deletion 69-70, deletion 144, N501Y, A570D, D614G, P681H, T716I, S982A, and D1118H.

B.1.351 (501.V2)

B.1.351 lineage and colloquially known as South African COVID-19 variant, is a variant of SARS-CoV-2. Preliminary results indicate that this variant may have an increased transmissibility. The B.1.351 variant is defined by multiple spike protein changes including: L18F, D80A, D215G, deletion 242-244, R246I, K417N, E484K, N501Y, D614G and A701V. There are three mutations of particular interest in the spike region of the B.1.351 genome: K417N, E484K, N501Y.

B.1.1.298 (Cluster 5)

B.1.1.298 was discovered in North Jutland, Denmark, and is believed to have been spread from minks to humans via mink farms. Several different mutations in the spike protein of the virus have been confirmed. The specific mutations include deletion 69–70, Y453F, D614G, I692V, M1229I, and optionally S1147L.

P.1 (B.1.1.248)

Lineage B.1.1.248, known as the Brazil(ian) variant, is one of the variants of SARS-CoV-2 which has been named P.1 lineage. P.1 has a number of S-protein modifications [L18F, T20N, P26S, D138Y, R190S, K417T, E484K, N501Y, D614G, H655Y, T1027I, V1176F] and is similar in certain key RBD positions (K417, E484, N501) to variant B.1.351 from South Africa.

B.1.427/B.1.429 (CAL.20C)

Lineage B.1.427/B.1.429, also known as CAL.20C, is defined by the following modifications in the S-protein: S13I, W152C, L452R, and D614G of which the L452R modification is of particular concern. CDC has listed B.1.427/B.1.429 as "variant of concern".

B.1.525

CA 03215771 2023-10-17

B.1.525 carries the same E484K modification as found in the P.1, and B.1.351 variants, and also carries the same ΔH69/ΔV70 deletion as found in B.1.1.7, and B.1.1.298. It also carries the modifications D614G, Q677H and F888L.

B.1.526

B.1.526 was detected as an emerging lineage of viral isolates in the New York region that shares mutations with previously reported variants. The most common sets of spike mutations in this lineage are L5F, T95I, D253G, E484K, D614G, and A701V.

The following table shows an overview of circulating SARS-CoV-2 strains which are VOI/VOC.

Table 1: Overview of circulating SARS-CoV-2 strains which are VOI/VOC

| Lineage | Amino acid substitution | | |

In one embodiment, the vaccine antigen described herein comprises, consists essentially of or consists of an amino acid sequence that is a modified variant of a spike protein (S) of SARS-CoV-2, or a fragment thereof.

In one embodiment, a vaccine antigen comprises an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 17 to 1273 of SEQ ID NO: 1 or 7, or an immunogenic fragment of the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 17 to 1273 of SEQ ID NO: 1 or 7.

In one embodiment, RNA encoding a vaccine antigen (i) comprises a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of nucleotides 49 to 3819 of SEQ ID NO: 2, 8 or 9, or a fragment of the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of nucleotides 49 to 3819 of SEQ ID NO: 2, 8 or 9; and/or (ii) encodes an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 17 to 1273 of SEQ ID NO: 1 or 7, or an immunogenic fragment of the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 17 to 1273 of SEQ ID NO: 1 or 7.

In one embodiment, the vaccine antigen comprises, consists essentially of or consists of an amino acid sequence that is a modified variant of a SARS-CoV-2 spike S1 fragment (S1) (the S1 subunit of a spike protein (S) of SARS-CoV-2), or a fragment thereof.

In one embodiment, a vaccine antigen comprises an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 17 to 683 of SEQ ID NO: 1, or an immunogenic fragment of the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 17 to 683 of SEQ ID NO: 1.

In one embodiment, RNA encoding a vaccine antigen (i) comprises a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of nucleotides 49 to 2049 of SEQ ID NO: 2, 8 or 9, or a fragment of the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of nucleotides 49 to 2049 of SEQ ID NO: 2, 8 or 9; and/or (ii) encodes an

CA 03215771 2023-10-17

amino acid sequence comprising an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 17 to 683 of SEQ ID NO: 1, or an immunogenic fragment of the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 17 to 683 of SEQ ID NO: 1.

In one embodiment, a vaccine antigen comprises an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 17 to 685 of SEQ ID NO: 1, or an immunogenic fragment of the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 17 to 685 of SEQ ID NO: 1.

In one embodiment, RNA encoding a vaccine antigen (i) comprises a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of nucleotides 49 to 2055 of SEQ ID NO: 2, 8 or 9, or a fragment of the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of nucleotides 49 to 2055 of SEQ ID NO: 2, 8 or 9; and/or (ii) encodes an amino acid sequence comprising an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 17 to 685 of SEQ ID NO: 1, or an immunogenic fragment of the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 17 to 685 of SEQ ID NO: 1.

In one embodiment, the vaccine antigen comprises, consists essentially of or consists of an amino acid sequence that is a modified variant of the receptor binding domain (RBD) of the S1 subunit of a spike protein (S) of SARS-CoV-2, or a fragment thereof. The amino acid sequence of amino acids 327 to 528 of SEQ ID NO: 1, or a fragment thereof is also referred to herein as "RBD" or "RBD domain".

In one embodiment, a vaccine antigen comprises an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 327 to 528 of SEQ ID NO: 1, or an immunogenic fragment of the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 327 to 528 of SEQ ID NO: 1.

In one embodiment, RNA encoding a vaccine antigen (i) comprises a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of nucleotides 979 to 1584 of SEQ ID NO: 2, 8 or 9, or a fragment of the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of nucleotides 979 to 1584 of SEQ ID NO: 2, 8 or 9; and/or (ii) encodes an amino acid sequence comprising an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 327 to 528 of SEQ ID NO: 1, or an immunogenic fragment of the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 327 to 528 of SEQ ID NO: 1.

According to certain embodiments, a signal peptide is fused, either directly or through a linker, to an amino acid sequence that is a modified variant of a SARS-CoV-2 S protein, or a fragment thereof, i.e., the antigenic peptide or protein. Accordingly, in one embodiment, a signal peptide is fused to the above described amino acid sequences derived from SARS-CoV-2 S protein or immunogenic fragments thereof (antigenic peptides or proteins) comprised by the vaccine antigens described above.

Such signal peptides are sequences, which typically exhibit a length of about 15 to 30 amino acids and are preferably located at the N-terminus of the antigenic peptide or protein, without being limited thereto. Signal peptides as defined herein preferably allow the transport of the antigenic peptide or protein as encoded by the RNA into a defined cellular compartment, preferably the cell surface, the endoplasmic reticulum (ER) or the endosomal-lysosomal compartment. In one embodiment, the signal peptide sequence as defined herein includes, without being limited thereto, the signal peptide sequence of SARS-CoV-2 S protein, in particular a sequence comprising the amino acid sequence of amino acids 1 to 16 or 1 to 19 of SEQ ID NO: 1 or a functional variant thereof.

In one embodiment, a signal sequence comprises the amino acid sequence of amino acids 1 to 16 of SEQ ID NO: 1, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 1 to 16 of SEQ ID NO: 1, or a functional fragment of the amino acid sequence of amino acids 1 to 16 of SEQ ID NO: 1, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to

CA 03215771 2023-10-17

the amino acid sequence of amino acids 1 to 16 of SEQ ID NO: 1. In one embodiment, a signal sequence comprises the amino acid sequence of amino acids 1 to 16 of SEQ ID NO: 1.

In one embodiment, RNA encoding a signal sequence (i) comprises the nucleotide sequence of nucleotides 1 to 48 of SEQ ID NO: 2, 8 or 9, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of nucleotides 1 to 48 of SEQ ID NO: 2, 8 or 9, or a fragment of the nucleotide sequence of nucleotides 1 to 48 of SEQ ID NO: 2, 8 or 9, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of nucleotides 1 to 48 of SEQ ID NO: 2, 8 or 9; and/or (ii) encodes an amino acid sequence comprising the amino acid sequence of amino acids 1 to 16 of SEQ ID NO: 1, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 1 to 16 of SEQ ID NO: 1, or a functional fragment of the amino acid sequence of amino acids 1 to 16 of SEQ ID NO: 1, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 1 to 16 of SEQ ID NO: 1. In one embodiment, RNA encoding a signal sequence (i) comprises the nucleotide sequence of nucleotides 1 to 48 of SEQ ID NO: 2, 8 or 9; and/or (ii) encodes an amino acid sequence comprising the amino acid sequence of amino acids 1 to 16 of SEQ ID NO: 1.

In one embodiment, a signal sequence comprises the amino acid sequence of amino acids 1 to 19 of SEQ ID NO: 1, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 1 to 19 of SEQ ID NO: 1, or a functional fragment of the amino acid sequence of amino acids 1 to 19 of SEQ ID NO: 1, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 1 to 19 of SEQ ID NO: 1. In one embodiment, a signal sequence comprises the amino acid sequence of amino acids 1 to 19 of SEQ ID NO: 1.

In one embodiment, RNA encoding a signal sequence (i) comprises the nucleotide sequence of nucleotides 1 to 57 of SEQ ID NO: 2, 8 or 9, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of nucleotides 1 to 57 of SEQ ID NO: 2, 8 or 9, or a fragment of the nucleotide sequence of nucleotides 1 to 57 of SEQ ID NO: 2, 8 or 9, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of nucleotides 1 to 57 of SEQ ID NO: 2, 8 or

9; and/or (ii) encodes an amino acid sequence comprising the amino acid sequence of amino acids 1 to 19 of SEQ ID NO: 1, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 1 to 19 of SEQ ID NO: 1, or a functional fragment of the amino acid sequence of amino acids 1 to 19 of SEQ ID NO: 1, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 1 to 19 of SEQ ID NO: 1. In one embodiment, RNA encoding a signal sequence (i) comprises the nucleotide sequence of nucleotides 1 to 57 of SEQ ID NO: 2, 8 or 9; and/or (ii) encodes an amino acid sequence comprising the amino acid sequence of amino acids 1 to 19 of SEQ ID NO: 1.

The signal peptide sequence as defined herein further includes, without being limited thereto, the signal peptide sequence of an immunoglobulin, e.g., the signal peptide sequence of an immunoglobulin heavy chain variable region, wherein the immunoglobulin may be human immunoglobulin. In particular, the signal peptide sequence as defined herein includes a sequence comprising the amino acid sequence of amino acids 1 to 22 of SEQ ID NO: 31 or a functional variant thereof.

In one embodiment, a signal sequence comprises the amino acid sequence of amino acids 1 to 22 of SEQ ID NO: 31, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 1 to 22 of SEQ ID NO: 31, or a functional fragment of the amino acid sequence of amino acids 1 to 22 of SEQ ID NO: 31, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 1 to 22 of SEQ ID NO: 31. In one embodiment, a signal sequence comprises the amino acid sequence of amino acids 1 to 22 of SEQ ID NO: 31.

In one embodiment, RNA encoding a signal sequence (i) comprises the nucleotide sequence of nucleotides 54 to 119 of SEQ ID NO: 32, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of nucleotides 54 to 119 of SEQ ID NO: 32, or a fragment of the nucleotide sequence of nucleotides 54 to 119 of SEQ ID NO: 32, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of nucleotides 54 to 119 of SEQ ID NO: 32; and/or (ii) encodes an amino acid sequence comprising the amino acid sequence of amino

CA 03215771 2023-10-17

acids 1 to 22 of SEQ ID NO: 31, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 1 to 22 of SEQ ID NO: 31, or a functional fragment of the amino acid sequence of amino acids 1 to 22 of SEQ ID NO: 31, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 1 to 22 of SEQ ID NO: 31. In one embodiment, RNA encoding a signal sequence (i) comprises the nucleotide sequence of nucleotides 54 to 119 of SEQ ID NO: 32; and/or (ii) encodes an amino acid sequence comprising the amino acid sequence of amino acids 1 to 22 of SEQ ID NO: 31.

Such signal peptides are preferably used in order to promote secretion of the encoded antigenic peptide or protein. More preferably, a signal peptide as defined herein is fused to an encoded antigenic peptide or protein as defined herein.

Accordingly, in particularly preferred embodiments, the RNA described herein comprises at least one coding region encoding an antigenic peptide or protein and a signal peptide, said signal peptide preferably being fused to the antigenic peptide or protein, more preferably to the N-terminus of the antigenic peptide or protein as described herein.

In one embodiment, a vaccine antigen comprises an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 1 or 7, or an immunogenic fragment of the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 1 or 7.

In one embodiment, RNA encoding a vaccine antigen (i) comprises a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 2, 8 or 9, or a fragment of the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 2, 8 or 9; and/or (ii) encodes an amino acid sequence comprising an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 1 or 7, or an immunogenic fragment of the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 1 or 7.

In one embodiment, a vaccine antigen comprises an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 7,

CA 03215771 2023-10-17

or an immunogenic fragment of the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 7.

In one embodiment, RNA encoding a vaccine antigen (i) comprises a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 15, 16, 19, 20, 24, or 25, or a fragment of the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 15, 16, 19, 20, 24, or 25; and/or (ii) encodes an amino acid sequence comprising an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 7, or an immunogenic fragment of the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 7.

In one embodiment, a vaccine antigen comprises an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 1 to 683 of SEQ ID NO: 1, or an immunogenic fragment of the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 1 to 683 of SEQ ID NO: 1.

In one embodiment, RNA encoding a vaccine antigen (i) comprises a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of nucleotides 1 to 2049 of SEQ ID NO: 2, 8 or 9, or a fragment of the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of nucleotides 1 to 2049 of SEQ ID NO: 2, 8 or 9; and/or (ii) encodes an amino acid sequence comprising an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 1 to 683 of SEQ ID NO: 1, or an immunogenic fragment of the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 1 to 683 of SEQ ID NO: 1.

In one embodiment, a vaccine antigen comprises an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 1 to 685 of SEQ ID NO: 1, or an immunogenic fragment of the amino acid sequence having at

CA 03215771 2023-10-17

least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 1 to 685 of SEQ ID NO: 1.

In one embodiment, RNA encoding a vaccine antigen (i) comprises a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of nucleotides 1 to 2055 of SEQ ID NO: 2, 8 or 9, or a fragment of the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of nucleotides 1 to 2055 of SEQ ID NO: 2, 8 or 9; and/or (ii) encodes an amino acid sequence comprising an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 1 to 685 of SEQ ID NO: 1, or an immunogenic fragment of the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 1 to 685 of SEQ ID NO: 1.

In one embodiment, a vaccine antigen comprises an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 3, or an immunogenic fragment of the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 3.

In one embodiment, RNA encoding a vaccine antigen (i) comprises a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 4, or a fragment of the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 4; and/or (ii) encodes an amino acid sequence comprising an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 3, or an immunogenic fragment of the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 3.

In one embodiment, a vaccine antigen comprises an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 1 to 221 of SEQ ID NO: 29, or an immunogenic fragment of the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 1 to 221 of SEQ ID NO: 29.

CA 03215771 2023-10-17

In one embodiment, RNA encoding a vaccine antigen (i) comprises a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of nucleotides 54 to 716 of SEQ ID NO: 30, or a fragment of the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of nucleotides 54 to 716 of SEQ ID NO: 30; and/or (ii) encodes an amino acid sequence comprising an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 1 to 221 of SEQ ID NO: 29, or an immunogenic fragment of the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 1 to 221 of SEQ ID NO: 29.

In one embodiment, a vaccine antigen comprises an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 1 to 224 of SEQ ID NO: 31, or an immunogenic fragment of the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 1 to 224 of SEQ ID NO: 31.

In one embodiment, RNA encoding a vaccine antigen (i) comprises a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of nucleotides 54 to 725 of SEQ ID NO: 32, or a fragment of the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of nucleotides 54 to 725 of SEQ ID NO: 32; and/or (ii) encodes an amino acid sequence comprising an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 1 to 224 of SEQ ID NO: 31, or an immunogenic fragment of the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 1 to 224 of SEQ ID NO: 31.

According to certain embodiments, a trimerization domain is fused, either directly or through a linker, e.g., a glycine/serine linker, to an amino acid sequence that is a modified variant of a SARS-CoV-2 S protein, or a fragment thereof, i.e., the antigenic peptide or protein. Accordingly, in one embodiment, a trimerization domain is fused to the above described amino acid sequences derived from SARS-CoV-2 S protein or immunogenic fragments thereof

CA 03215771 2023-10-17

(antigenic peptides or proteins) comprised by the vaccine antigens described above (which may optionally be fused to a signal peptide as described above).

Such trimerization domains are preferably located at the C-terminus of the antigenic peptide or protein, without being limited thereto. Trimerization domains as defined herein preferably allow the trimerization of the antigenic peptide or protein as encoded by the RNA. Examples of trimerization domains as defined herein include, without being limited thereto, foldon, the natural trimerization domain of T4 fibritin. The C-terminal domain of T4 fibritin (foldon) is obligatory for the formation of the fibritin trimer structure and can be used as an artificial trimerization domain. In one embodiment, the trimerization domain as defined herein includes, without being limited thereto, a sequence comprising the amino acid sequence of amino acids 3 to 29 of SEQ ID NO: 10 or a functional variant thereof. In one embodiment, the trimerization domain as defined herein includes, without being limited thereto, a sequence comprising the amino acid sequence of SEQ ID NO: 10 or a functional variant thereof.

In one embodiment, a trimerization domain comprises the amino acid sequence of amino acids 3 to 29 of SEQ ID NO: 10, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 3 to 29 of SEQ ID NO: 10, or a functional fragment of the amino acid sequence of amino acids 3 to 29 of SEQ ID NO: 10, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 3 to 29 of SEQ ID NO: 10. In one embodiment, a trimerization domain comprises the amino acid sequence of amino acids 3 to 29 of SEQ ID NO: 10.

In one embodiment, RNA encoding a trimerization domain (i) comprises the nucleotide sequence of nucleotides 7 to 87 of SEQ ID NO: 11, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of nucleotides 7 to 87 of SEQ ID NO: 11, or a fragment of the nucleotide sequence of nucleotides 7 to 87 of SEQ ID NO: 11, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of nucleotides 7 to 87 of SEQ ID NO: 11; and/or (ii) encodes an amino acid sequence comprising the amino acid sequence of amino acids 3 to 29 of SEQ ID NO: 10, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 3 to 29 of SEQ ID

CA 03215771 2023-10-17

NO: 10, or a functional fragment of the amino acid sequence of amino acids 3 to 29 of SEQ ID NO: 10, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 3 to 29 of SEQ ID NO: 10. In one embodiment, RNA encoding a trimerization domain (i) comprises the nucleotide sequence of nucleotides 7 to 87 of SEQ ID NO: 11; and/or (ii) encodes an amino acid sequence comprising the amino acid sequence of amino acids 3 to 29 of SEQ ID NO: 10.

In one embodiment, a trimerization domain comprises the amino acid sequence SEQ ID NO: 10, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 10, or a functional fragment of the amino acid sequence of SEQ ID NO: 10, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 10. In one embodiment, a trimerization domain comprises the amino acid sequence of SEQ ID NO: 10.

In one embodiment, RNA encoding a trimerization domain (i) comprises the nucleotide sequence of SEQ ID NO: 11, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 11, or a fragment of the nucleotide sequence of SEQ ID NO: 11, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 11; and/or (ii) encodes an amino acid sequence comprising the amino acid sequence of SEQ ID NO: 10, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 10, or a functional fragment of the amino acid sequence of SEQ ID NO: 10, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 10. In one embodiment, RNA encoding a trimerization domain (i) comprises the nucleotide sequence of SEQ ID NO: 11; and/or (ii) encodes an amino acid sequence comprising the amino acid sequence of SEQ ID NO: 10.

Such trimerization domains are preferably used in order to promote trimerization of the encoded antigenic peptide or protein. More preferably, a trimerization domain as defined herein is fused to an antigenic peptide or protein as defined herein.

Accordingly, in particularly preferred embodiments, the RNA described herein comprises at least one coding region encoding an antigenic peptide or protein and a trimerization domain

as defined herein, said trimerization domain preferably being fused to the antigenic peptide or protein, more preferably to the C-terminus of the antigenic peptide or protein.

In one embodiment, a vaccine antigen comprises an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 5, or an immunogenic fragment of the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 5.

In one embodiment, RNA encoding a vaccine antigen (i) comprises a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 6, or a fragment of the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 6; and/or (ii) encodes an amino acid sequence comprising an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 5, or an immunogenic fragment of the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 5.

In one embodiment, RNA encoding a vaccine antigen (i) comprises a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 17, 21, or 26, or a fragment of the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 17, 21, or 26; and/or (ii) encodes an amino acid sequence comprising an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 5, or an immunogenic fragment of the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 5.

In one embodiment, a vaccine antigen comprises an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 18, or an immunogenic fragment of the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 18.

In one embodiment, a vaccine antigen comprises an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 1 to 257 of SEQ ID NO: 29, or an immunogenic fragment of the amino acid sequence having at

CA 03215771 2023-10-17

least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 1 to 257 of SEQ ID NO: 29.

In one embodiment, RNA encoding a vaccine antigen (i) comprises a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of nucleotides 54 to 824 of SEQ ID NO: 30, or a fragment of the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of nucleotides 54 to 824 of SEQ ID NO: 30; and/or (ii) encodes an amino acid sequence comprising an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 1 to 257 of SEQ ID NO: 29, or an immunogenic fragment of the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 1 to 257 of SEQ ID NO: 29.

In one embodiment, a vaccine antigen comprises an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 1 to 260 of SEQ ID NO: 31, or an immunogenic fragment of the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 1 to 260 of SEQ ID NO: 31.

In one embodiment, RNA encoding a vaccine antigen (i) comprises a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of nucleotides 54 to 833 of SEQ ID NO: 32, or a fragment of the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of nucleotides 54 to 833 of SEQ ID NO: 32; and/or (ii) encodes an amino acid sequence comprising an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 1 to 260 of SEQ ID NO: 31, or an immunogenic fragment of the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 1 to 260 of SEQ ID NO: 31.

In one embodiment, a vaccine antigen comprises an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 20 to 257 of SEQ ID NO: 29, or an immunogenic fragment of the amino acid sequence having at

CA 03215771 2023-10-17

least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 20 to 257 of SEQ ID NO: 29.

In one embodiment, RNA encoding a vaccine antigen (i) comprises a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of nucleotides 111 to 824 of SEQ ID NO: 30, or a fragment of the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of nucleotides 111 to 824 of SEQ ID NO: 30; and/or (ii) encodes an amino acid sequence comprising an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 20 to 257 of SEQ ID NO: 29, or an immunogenic fragment of the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 20 to 257 of SEQ ID NO: 29.

In one embodiment, a vaccine antigen comprises an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 23 to 260 of SEQ ID NO: 31, or an immunogenic fragment of the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 23 to 260 of SEQ ID NO: 31.

In one embodiment, RNA encoding a vaccine antigen (i) comprises a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of nucleotides 120 to 833 of SEQ ID NO: 32, or a fragment of the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of nucleotides 120 to 833 of SEQ ID NO: 32; and/or (ii) encodes an amino acid sequence comprising an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 23 to 260 of SEQ ID NO: 31, or an immunogenic fragment of the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 23 to 260 of SEQ ID NO: 31.

According to certain embodiments, a transmembrane domain domain is fused, either directly or through a linker, e.g., a glycine/serine linker, to an amino acid sequence that is a modified variant of a SARS-CoV-2 S protein, or a fragment thereof, i.e., the antigenic peptide or protein.

CA 03215771 2023-10-17

Accordingly, in one embodiment, a transmembrane domain is fused to the above described amino acid sequences derived from SARS-CoV-2 S protein or immunogenic fragments thereof (antigenic peptides or proteins) comprised by the vaccine antigens described above (which may optionally be fused to a signal peptide and/or trimerization domain as described above). Such transmembrane domains are preferably located at the C-terminus of the antigenic peptide or protein, without being limited thereto. Preferably, such transmembrane domains are located at the C-terminus of the trimerization domain, if present, without being limited thereto. In one embodiment, a trimerization domain is present between an amino acid sequence that is a modified variant of a SARS-CoV-2 S protein, or a fragment thereof, i.e., the antigenic peptide or protein, and the transmembrane domain.

Transmembrane domains as defined herein preferably allow the anchoring into a cellular membrane of the antigenic peptide or protein.

In one embodiment, the transmembrane domain sequence as defined herein includes, without being limited thereto, the transmembrane domain sequence of SARS-CoV-2 S protein, in particular a sequence comprising the amino acid sequence of amino acids 1207 to 1254 of SEQ ID NO: 1 or a functional variant thereof.

In one embodiment, a transmembrane domain sequence comprises the amino acid sequence of amino acids 1207 to 1254 of SEQ ID NO: 1, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 1207 to 1254 of SEQ ID NO: 1, or a functional fragment of the amino acid sequence of amino acids 1207 to 1254 of SEQ ID NO: 1, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 1207 to 1254 of SEQ ID NO: 1. In one embodiment, a transmembrane domain sequence comprises the amino acid sequence of amino acids 1207 to 1254 of SEQ ID NO: 1.

In one embodiment, RNA encoding a transmembrane domain sequence (i) comprises the nucleotide sequence of nucleotides 3619 to 3762 of SEQ ID NO: 2, 8 or 9, a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of nucleotides 3619 to 3762 of SEQ ID NO: 2, 8 or 9, or a fragment of the nucleotide sequence of nucleotides 3619 to 3762 of SEQ ID NO: 2, 8 or 9, or the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the

CA 03215771 2023-10-17

nucleotide sequence of nucleotides 3619 to 3762 of SEQ ID NO: 2, 8 or 9; and/or (ii) encodes an amino acid sequence comprising the amino acid sequence of amino acids 1207 to 1254 of SEQ ID NO: 1, an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 1207 to 1254 of SEQ ID NO: 1, or a functional fragment of the amino acid sequence of amino acids 1207 to 1254 of SEQ ID NO: 1, or the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 1207 to 1254 of SEQ ID NO: 1. In one embodiment, RNA encoding a transmembrane domain sequence (i) comprises the nucleotide sequence of nucleotides 3619 to 3762 of SEQ ID NO: 2, 8 or 9; and/or (ii) encodes an amino acid sequence comprising the amino acid sequence of amino acids 1207 to 1254 of SEQ ID NO: 1.

In one embodiment, a vaccine antigen comprises an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 1 to 311 of SEQ ID NO: 29, or an immunogenic fragment of the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 1 to 311 of SEQ ID NO: 29.

In one embodiment, RNA encoding a vaccine antigen (i) comprises a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of nucleotides 54 to 986 of SEQ ID NO: 30, or a fragment of the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of nucleotides 54 to 986 of SEQ ID NO: 30; and/or (ii) encodes an amino acid sequence comprising an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 1 to 311 of SEQ ID NO: 29, or an immunogenic fragment of the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 1 to 311 of SEQ ID NO: 29.

In one embodiment, a vaccine antigen comprises an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 1 to 314 of SEQ ID NO: 31, or an immunogenic fragment of the amino acid sequence having at

CA 03215771 2023-10-17

least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 1 to 314 of SEQ ID NO: 31.

In one embodiment, RNA encoding a vaccine antigen (i) comprises a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of nucleotides 54 to 995 of SEQ ID NO: 32, or a fragment of the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of nucleotides 54 to 995 of SEQ ID NO: 32; and/or (ii) encodes an amino acid sequence comprising an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 1 to 314 of SEQ ID NO: 31, or an immunogenic fragment of the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 1 to 314 of SEQ ID NO: 31.

In one embodiment, a vaccine antigen comprises an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 20 to 311 of SEQ ID NO: 29, or an immunogenic fragment of the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 20 to 311 of SEQ ID NO: 29.

In one embodiment, RNA encoding a vaccine antigen (i) comprises a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of nucleotides 111 to 986 of SEQ ID NO: 30, or a fragment of the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of nucleotides 111 to 986 of SEQ ID NO: 30; and/or (ii) encodes an amino acid sequence comprising an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 20 to 311 of SEQ ID NO: 29, or an immunogenic fragment of the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 20 to 311 of SEQ ID NO: 29.

In one embodiment, a vaccine antigen comprises an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 23 to 314 of SEQ ID NO: 31, or an immunogenic fragment of the amino acid sequence having at

least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 23 to 314 of SEQ ID NO: 31.

In one embodiment, RNA encoding a vaccine antigen (i) comprises a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of nucleotides 120 to 995 of SEQ ID NO: 32, or a fragment of the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of nucleotides 120 to 995 of SEQ ID NO: 32; and/or (ii) encodes an amino acid sequence comprising an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 23 to 314 of SEQ ID NO: 31, or an immunogenic fragment of the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 23 to 314 of SEQ ID NO: 31.

In one embodiment, RNA encoding a vaccine antigen (i) comprises a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 30, or a fragment of the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 30; and/or (ii) encodes an amino acid sequence comprising an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 29, or an immunogenic fragment of the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 29.

In one embodiment, RNA encoding a vaccine antigen (i) comprises a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 32, or a fragment of the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 32; and/or (ii) encodes an amino acid sequence comprising an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 31, or an immunogenic fragment of the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 31.

In one embodiment, a vaccine antigen comprises an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 28,

CA 03215771 2023-10-17

or an immunogenic fragment of the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 28.

In one embodiment, RNA encoding a vaccine antigen (i) comprises a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 27, or a fragment of the nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 27; and/or (ii) encodes an amino acid sequence comprising an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 28, or an immunogenic fragment of the amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 28.

In one embodiment, the vaccine antigens described above comprise an amino acid sequence that is a modified variant of a contiguous sequence of SARS-CoV-2 coronavirus spike (S) protein that consists of or essentially consists of the above described amino acid sequences derived from SARS-CoV-2 S protein or immunogenic fragments thereof (antigenic peptides or proteins) comprised by the vaccine antigens described above. In one embodiment, the vaccine antigens described above comprise an amino acid sequence that is a modified variant of a contiguous sequence of SARS-CoV-2 coronavirus spike (S) protein of no more than 220 amino acids, 215 amino acids, 210 amino acids, or 205 amino acids.

In one embodiment, RNA encoding a vaccine antigen is nucleoside modified messenger RNA (modRNA) that is a modified variant of an RNA described herein as BNT162b1 (RBP020.3), BNT162b2 (RBP020.1 or RBP020.2). In one embodiment, RNA encoding a vaccine antigen is nucleoside modified messenger RNA (modRNA) that is a modified variant of an RNA described herein as RBP020.2.

In one embodiment, RNA encoding a vaccine antigen is nucleoside modified messenger RNA (modRNA) and (i) comprises a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 21, and/or (ii) encodes an amino acid sequence comprising an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 5.

In one embodiment, RNA encoding a vaccine antigen is nucleoside modified messenger RNA (modRNA) and (i) comprises a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%,

CA 03215771 2023-10-17

90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 19, or 20, and/or (ii) encodes an amino acid sequence comprising an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 7.

In one embodiment, RNA encoding a vaccine antigen is nucleoside modified messenger RNA (modRNA) and (i) comprises a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 20, and/or (ii) encodes an amino acid sequence comprising an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of SEQ ID NO: 7.

As used herein, the term "vaccine" refers to a composition that induces an immune response upon inoculation into a subject. In some embodiments, the induced immune response provides protective immunity.

In one embodiment, the RNA encoding the antigen molecule is expressed in cells of the subject to provide the antigen molecule. In one embodiment, expression of the antigen molecule is at the cell surface or into the extracellular space. In one embodiment, the antigen molecule is presented in the context of MHC. In one embodiment, the RNA encoding the antigen molecule is transiently expressed in cells of the subject. In one embodiment, after administration of the RNA encoding the antigen molecule, in particular after intramuscular administration of the RNA encoding the antigen molecule, expression of the RNA encoding the antigen molecule in muscle occurs. In one embodiment, after administration of the RNA encoding the antigen molecule, expression of the RNA encoding the antigen molecule in spleen occurs. In one embodiment, after administration of the RNA encoding the antigen molecule, expression of the RNA encoding the antigen molecule in antigen presenting cells, preferably professional antigen presenting cells occurs. In one embodiment, the antigen presenting cells are selected from the group consisting of dendritic cells, macrophages and B cells. In one embodiment, after administration of the RNA encoding the antigen molecule, no or essentially no expression of the RNA encoding the antigen molecule in lung and/or liver occurs. In one embodiment, after administration of the RNA encoding the antigen molecule, expression of the RNA encoding the antigen molecule in spleen is at least 5-fold the amount of expression in lung.

In some embodiments, the methods and agents, e.g., mRNA compositions, described herein following administration, in particular following intramuscular administration, to a subject result in delivery of the RNA encoding a vaccine antigen to lymph nodes and/or spleen. In some embodiments, RNA encoding a vaccine antigen is detectable in lymph nodes and/or spleen 6 hours or later following administration and preferably up to 6 days or longer.

In some embodiments, the methods and agents, e.g., mRNA compositions, described herein following administration, in particular following intramuscular administration, to a subject result in delivery of the RNA encoding a vaccine antigen to B cell follicles, subcapsular sinus, and/or T cell zone, in particular B cell follicles and/or subcapsular sinus of lymph nodes.

In some embodiments, the methods and agents, e.g., mRNA compositions, described herein following administration, in particular following intramuscular administration, to a subject result in delivery of the RNA encoding a vaccine antigen to B cells (CD19+), subcapsular sinus macrophages (CD169+) and/or dendritic cells (CD11c+) in the T cell zone and intermediary sinus of lymph nodes, in particular to B cells (CD19+) and/or subcapsular sinus macrophages (CD169+) of lymph nodes.

In some embodiments, the methods and agents, e.g., mRNA compositions, described herein following administration, in particular following intramuscular administration, to a subject result in delivery of the RNA encoding a vaccine antigen to white pulp of spleen.

In some embodiments, the methods and agents, e.g., mRNA compositions, described herein following administration, in particular following intramuscular administration, to a subject result in delivery of the RNA encoding a vaccine antigen to B cells, DCs (CD11c+), in particular those surrounding the B cells, and/or mcrophages of spleen, in particular to B cells and/or DCs (CD11c+).

In one embodiment, the vaccine antigen is expressed in lymph node and/or spleen, in particular in the cells of lymph node and/or spleen described above.

An antigen molecule or a procession product thereof, e.g., a fragment thereof, may bind to an antigen receptor such as a BCR or TCR carried by immune effector cells, or to antibodies.

A peptide and protein antigen which is provided to a subject according to the invention, e.g., by administering RNA encoding the peptide and protein antigen, i.e., a vaccine antigen, preferably results in the induction of an immune response, e.g., a humoral and/or cellular

CA 03215771 2023-10-17

immune response in the subject being provided the peptide or protein antigen. Said immune response is preferably directed against a target antigen, in particular coronavirus S protein, in particular SARS-CoV-2 S protein.

A vaccine antigen may comprise the target antigen, a variant thereof, or a fragment thereof. In one embodiment, such fragment or variant is immunologically equivalent to the target antigen. The term "fragment of an antigen" or "variant of an antigen" means an agent which results in the induction of an immune response which immune response targets the antigen, i.e. a target antigen. Thus, a vaccine antigen may correspond to or may comprise the target antigen, may correspond to or may comprise a fragment of the target antigen or may correspond to or may comprise an antigen which is homologous to the target antigen or a fragment thereof. A vaccine antigen may comprise an immunogenic fragment of a target antigen or an amino acid sequence being homologous to an immunogenic fragment of a target antigen. An "immunogenic fragment of an antigen" preferably relates to a fragment of an antigen which is capable of inducing an immune response against the target antigen. The vaccine antigen may be a recombinant antigen.

The term "immunologically equivalent" means that the immunologically equivalent molecule such as the immunologically equivalent amino acid sequence exhibits the same or essentially the same immunological properties and/or exerts the same or essentially the same immunological effects, e.g., with respect to the type of the immunological effect. In the context of the present disclosure, the term "immunologically equivalent" is preferably used with respect to the immunological effects or properties of antigens or antigen variants used for immunization. For example, an amino acid sequence is immunologically equivalent to a reference amino acid sequence if said amino acid sequence when exposed to the immune system of a subject induces an immune reaction having a specificity of reacting with the reference amino acid sequence.

"Activation" or "stimulation", as used herein, refers to the state of an immune effector cell such as T cell that has been sufficiently stimulated to induce detectable cellular proliferation. Activation can also be associated with initiation of signaling pathways, induced cytokine production, and detectable effector functions. The term "activated immune effector cells" refers to, among other things, immune effector cells that are undergoing cell division.

CA 03215771 2023-10-17

The term "priming" refers to a process wherein an immune effector cell such as a T cell has its first contact with its specific antigen and causes differentiation into effector cells such as effector T cells.

The term "clonal expansion" or "expansion" refers to a process wherein a specific entity is multiplied. In the context of the present disclosure, the term is preferably used in the context of an immunological response in which immune effector cells are stimulated by an antigen, proliferate, and the specific immune effector cell recognizing said antigen is amplified. Preferably, clonal expansion leads to differentiation of the immune effector cells.

The term "antigen" relates to an agent comprising an epitope against which an immune response can be generated. The term "antigen" includes, in particular, proteins and peptides. In one embodiment, an antigen is presented by cells of the immune system such as antigen presenting cells like dendritic cells or macrophages. An antigen or a procession product thereof such as a T-cell epitope is in one embodiment bound by a T- or B-cell receptor, or by an immunoglobulin molecule such as an antibody. Accordingly, an antigen or a procession product thereof may react specifically with antibodies or T lymphocytes (T cells). In one embodiment, an antigen is a viral antigen, such as a coronavirus S protein, e.g., SARS-CoV-2 S protein, and an epitope is derived from such antigen.

The term "viral antigen" refers to any viral component having antigenic properties, i.e. being able to provoke an immune response in an individual. The viral antigen may be coronavirus S protein, e.g., SARS-CoV-2 S protein.

The term "expressed on the cell surface" or "associated with the cell surface" means that a molecule such as an antigen is associated with and located at the plasma membrane of a cell, wherein at least a part of the molecule faces the extracellular space of said cell and is accessible from the outside of said cell, e.g., by antibodies located outside the cell. In this context, a part is preferably at least 4, preferably at least 8, preferably at least 12, more preferably at least 20 amino acids. The association may be direct or indirect. For example, the association may be by one or more transmembrane domains, one or more lipid anchors, or by the interaction with any other protein, lipid, saccharide, or other structure that can be found on the outer leaflet of the plasma membrane of a cell. For example, a molecule associated with the surface of a cell may be a transmembrane protein having an extracellular portion or

CA 03215771 2023-10-17

may be a protein associated with the surface of a cell by interacting with another protein that is a transmembrane protein.

"Cell surface" or "surface of a cell" is used in accordance with its normal meaning in the art, and thus includes the outside of the cell which is accessible to binding by proteins and other molecules. An antigen is expressed on the surface of cells if it is located at the surface of said cells and is accessible to binding by e.g. antigen-specific antibodies added to the cells.

The term "extracellular portion" or "exodomain" in the context of the present invention refers to a part of a molecule such as a protein that is facing the extracellular space of a cell and preferably is accessible from the outside of said cell, e.g., by binding molecules such as antibodies located outside the cell. Preferably, the term refers to one or more extracellular loops or domains or a fragment thereof.

The term "epitope" refers to a part or fragment of a molecule such as an antigen that is recognized by the immune system. For example, the epitope may be recognized by T cells, B cells or antibodies. An epitope of an antigen may include a continuous or discontinuous portion of the antigen and may be between about 5 and about 100, such as between about 5 and about 50, more preferably between about 8 and about 30, most preferably between about 8 and about 25 amino acids in length, for example, the epitope may be preferably 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25 amino acids in length. In one embodiment, an epitope is between about 10 and about 25 amino acids in length. The term "epitope" includes T cell epitopes.

The term "T cell epitope" refers to a part or fragment of a protein that is recognized by a T cell when presented in the context of MHC molecules. The term "major histocompatibility complex" and the abbreviation "MHC" includes MHC class I and MHC class II molecules and relates to a complex of genes which is present in all vertebrates. MHC proteins or molecules are important for signaling between lymphocytes and antigen presenting cells or diseased cells in immune reactions, wherein the MHC proteins or molecules bind peptide epitopes and present them for recognition by T cell receptors on T cells. The proteins encoded by the MHC are expressed on the surface of cells, and display both self-antigens (peptide fragments from the cell itself) and non-self-antigens (e.g., fragments of invading microorganisms) to a T cell. In the case of class I MHC/peptide complexes, the binding peptides are typically about 8 to

CA 03215771 2023-10-17

about 10 amino acids long although longer or shorter peptides may be effective. In the case of class II MHC/peptide complexes, the binding peptides are typically about 10 to about 25 amino acids long and are in particular about 13 to about 18 amino acids long, whereas longer and shorter peptides may be effective.

The peptide and protein antigen can be 2-100 amino acids, including for example, 5 amino acids, 10 amino acids, 15 amino acids, 20 amino acids, 25 amino acids, 30 amino acids, 35 amino acids, 40 amino acids, 45 amino acids, or 50 amino acids in length. In some embodiments, a peptide can be greater than 50 amino acids. In some embodiments, the peptide can be greater than 100 amino acids.

The peptide or protein antigen can be any peptide or protein that can induce or increase the ability of the immune system to develop antibodies and T cell responses to the peptide or protein.

In one embodiment, vaccine antigen is recognized by an immune effector cell. Preferably, the vaccine antigen if recognized by an immune effector cell is able to induce in the presence of appropriate co-stimulatory signals, stimulation, priming and/or expansion of the immune effector cell carrying an antigen receptor recognizing the vaccine antigen. In the context of the embodiments of the present invention, the vaccine antigen is preferably presented or present on the surface of a cell, preferably an antigen presenting cell. In one embodiment, an antigen is presented by a diseased cell such as a virus-infected cell. In one embodiment, an antigen receptor is a TCR which binds to an epitope of an antigen presented in the context of MHC. In one embodiment, binding of a TCR when expressed by T cells and/or present on T cells to an antigen presented by cells such as antigen presenting cells results in stimulation, priming and/or expansion of said T cells. In one embodiment, binding of a TCR when expressed by T cells and/or present on T cells to an antigen presented on diseased cells results in cytolysis and/or apoptosis of the diseased cells, wherein said T cells preferably release cytotoxic factors, e.g. perforins and granzymes.

In one embodiment, an antigen receptor is an antibody or B cell receptor which binds to an epitope in an antigen. In one embodiment, an antibody or B cell receptor binds to native epitopes of an antigen.

CA 03215771 2023-10-17

Nucleic acids

The term "polynucleotide" or "nucleic acid", as used herein, is intended to include DNA and RNA such as genomic DNA, cDNA, mRNA, recombinantly produced and chemically synthesized molecules. A nucleic acid may be single-stranded or double-stranded. RNA includes *in vitro* transcribed RNA (IVT RNA) or synthetic RNA. According to the invention, a polynucleotide is preferably isolated.

Nucleic acids may be comprised in a vector. The term "vector" as used herein includes any vectors known to the skilled person including plasmid vectors, cosmid vectors, phage vectors such as lambda phage, viral vectors such as retroviral, adenoviral or baculoviral vectors, or artificial chromosome vectors such as bacterial artificial chromosomes (BAC), yeast artificial chromosomes (YAC), or P1 artificial chromosomes (PAC). Said vectors include expression as well as cloning vectors. Expression vectors comprise plasmids as well as viral vectors and generally contain a desired coding sequence and appropriate DNA sequences necessary for the expression of the operably linked coding sequence in a particular host organism (e.g., bacteria, yeast, plant, insect, or mammal) or in *in vitro* expression systems. Cloning vectors are generally used to engineer and amplify a certain desired DNA fragment and may lack functional sequences needed for expression of the desired DNA fragments.

In one embodiment of all aspects of the invention, the RNA encoding the vaccine antigen is expressed in cells such as antigen presenting cells of the subject treated to provide the vaccine antigen.

The nucleic acids described herein may be recombinant and/or isolated molecules.

In the present disclosure, the term "RNA" relates to a nucleic acid molecule which includes ribonucleotide residues. In preferred embodiments, the RNA contains all or a majority of ribonucleotide residues. As used herein, "ribonucleotide" refers to a nucleotide with a hydroxyl group at the 2'-position of a β-D-ribofuranosyl group. RNA encompasses without limitation, double stranded RNA, single stranded RNA, isolated RNA such as partially purified RNA, essentially pure RNA, synthetic RNA, recombinantly produced RNA, as well as modified RNA that differs from naturally occurring RNA by the addition, deletion, substitution and/or alteration of one or more nucleotides. Such alterations may refer to addition of non-nucleotide material to internal RNA nucleotides or to the end(s) of RNA. It is also contemplated

CA 03215771 2023-10-17

herein that nucleotides in RNA may be non-standard nucleotides, such as chemically synthesized nucleotides or deoxynucleotides. For the present disclosure, these altered RNAs are considered analogs of naturally-occurring RNA.

In certain embodiments of the present disclosure, the RNA is messenger RNA (mRNA) that relates to a RNA transcript which encodes a peptide or protein. As established in the art, mRNA generally contains a 5' untranslated region (5'-UTR), a peptide coding region and a 3' untranslated region (3'-UTR). In some embodiments, the RNA is produced by *in vitro* transcription or chemical synthesis. In one embodiment, the mRNA is produced by *in vitro* transcription using a DNA template where DNA refers to a nucleic acid that contains deoxyribonucleotides.

In one embodiment, RNA is *in vitro* transcribed RNA (IVT-RNA) and may be obtained by *in vitro* transcription of an appropriate DNA template. The promoter for controlling transcription can be any promoter for any RNA polymerase. A DNA template for *in vitro* transcription may be obtained by cloning of a nucleic acid, in particular cDNA, and introducing it into an appropriate vector for *in vitro* transcription. The cDNA may be obtained by reverse transcription of RNA.

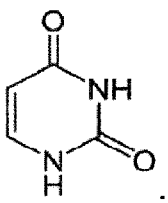
In certain embodiments of the present disclosure, the RNA is "replicon RNA" or simply a "replicon", in particular "self-replicating RNA" or "self-amplifying RNA". In one particularly preferred embodiment, the replicon or self-replicating RNA is derived from or comprises elements derived from a ssRNA virus, in particular a positive-stranded ssRNA virus such as an alphavirus. Alphaviruses are typical representatives of positive-stranded RNA viruses. Alphaviruses replicate in the cytoplasm of infected cells (for review of the alphaviral life cycle see José et al., Future Microbiol., 2009, vol. 4, pp. 837–856). The total genome length of many alphaviruses typically ranges between 11,000 and 12,000 nucleotides, and the genomic RNA typically has a 5'-cap, and a 3' poly(A) tail. The genome of alphaviruses encodes non-structural proteins (involved in transcription, modification and replication of viral RNA and in protein modification) and structural proteins (forming the virus particle). There are typically two open reading frames (ORFs) in the genome. The four non-structural proteins (nsP1–nsP4) are typically encoded together by a first ORF beginning near the 5' terminus of the genome, while alphavirus structural proteins are encoded together by a second ORF which is found downstream of the first ORF and extends near the 3' terminus of the genome. Typically, the

CA 03215771 2023-10-17

first ORF is larger than the second ORF, the ratio being roughly 2:1. In cells infected by an alphavirus, only the nucleic acid sequence encoding non-structural proteins is translated from the genomic RNA, while the genetic information encoding structural proteins is translatable from a subgenomic transcript, which is an RNA molecule that resembles eukaryotic messenger RNA (mRNA; Gould et al., 2010, Antiviral Res., vol. 87 pp. 111–124). Following infection, i.e. at early stages of the viral life cycle, the (+) stranded genomic RNA directly acts like a messenger RNA for the translation of the open reading frame encoding the non-structural poly-protein (nsP1234). Alphavirus-derived vectors have been proposed for delivery of foreign genetic information into target cells or target organisms. In simple approaches, the open reading frame encoding alphaviral structural proteins is replaced by an open reading frame encoding a protein of interest. Alphavirus-based trans-replication systems rely on alphavirus nucleotide sequence elements on two separate nucleic acid molecules: one nucleic acid molecule encodes a viral replicase, and the other nucleic acid molecule is capable of being replicated by said replicase in trans (hence the designation trans-replication system). Trans-replication requires the presence of both these nucleic acid molecules in a given host cell. The nucleic acid molecule capable of being replicated by the replicase in trans must comprise certain alphaviral sequence elements to allow recognition and RNA synthesis by the alphaviral replicase.

In one embodiment, the RNA described herein may have modified nucleosides. In some embodiments, the RNA comprises a modified nucleoside in place of at least one (e.g., every) uridine.

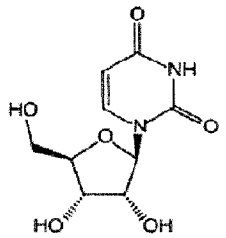
The term "uracil," as used herein, describes one of the nucleobases that can occur in the nucleic acid of RNA. The structure of uracil is:



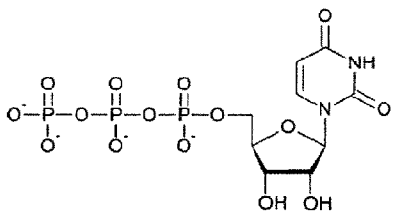
.

The term "uridine," as used herein, describes one of the nucleosides that can occur in RNA. The structure of uridine is:

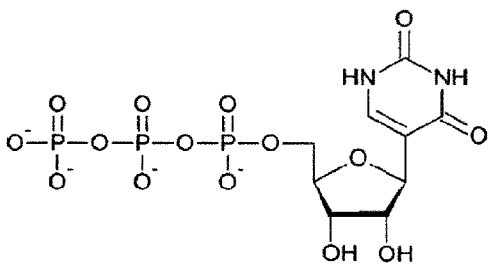
CA 03215771 2023-10-17



UTP (uridine 5'-triphosphate) has the following structure:

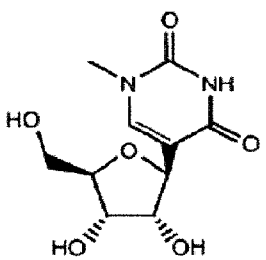


Pseudo-UTP (pseudouridine 5'-triphosphate) has the following structure:

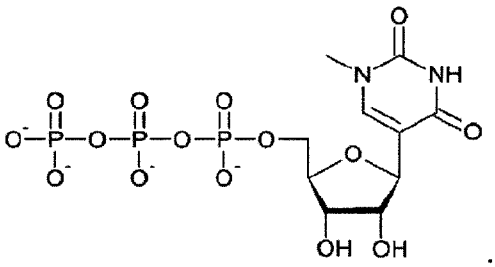


"Pseudouridine" is one example of a modified nucleoside that is an isomer of uridine, where the uracil is attached to the pentose ring via a carbon-carbon bond instead of a nitrogen-carbon glycosidic bond.

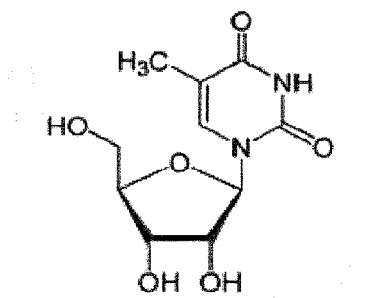
Another exemplary modified nucleoside is N1-methyl-pseudouridine (m1Ψ), which has the structure:



N1-methyl-pseudo-UTP has the following structure:



Another exemplary modified nucleoside is 5-methyl-uridine (m5U), which has the structure:



In some embodiments, one or more uridine in the RNA described herein is replaced by a modified nucleoside. In some embodiments, the modified nucleoside is a modified uridine. In some embodiments, RNA comprises a modified nucleoside in place of at least one uridine. In some embodiments, RNA comprises a modified nucleoside in place of each uridine.

In some embodiments, the modified nucleoside is independently selected from pseudouridine (ψ), N1-methyl-pseudouridine (m1ψ), and 5-methyl-uridine (m5U). In some embodiments, the modified nucleoside comprises pseudouridine (ψ). In some embodiments, the modified nucleoside comprises N1-methyl-pseudouridine (m1ψ). In some embodiments, the modified nucleoside comprises 5-methyl-uridine (m5U). In some embodiments, RNA may comprise more than one type of modified nucleoside, and the modified nucleosides are independently selected from pseudouridine (ψ), N1-methyl-pseudouridine (m1ψ), and 5-methyl-uridine (m5U). In some embodiments, the modified nucleosides comprise pseudouridine (ψ) and N1-methyl-pseudouridine (m1ψ). In some embodiments, the modified nucleosides comprise pseudouridine (ψ) and 5-methyl-uridine (m5U). In some embodiments, the modified nucleosides comprise N1-methyl-pseudouridine (m1ψ) and 5-methyl-uridine (m5U). In some embodiments, the modified nucleosides comprise pseudouridine (ψ), N1-methyl-pseudouridine (m1ψ), and 5-methyl-uridine (m5U).

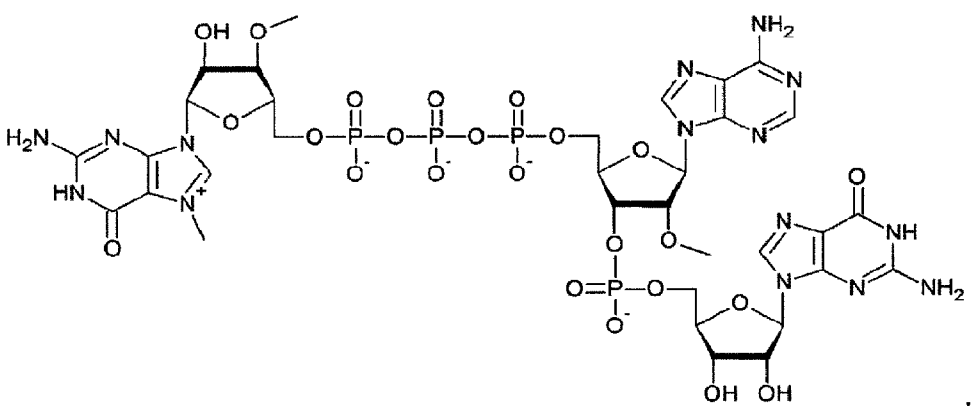
In some embodiments, the modified nucleoside replacing one or more, e.g., all, uridine in the RNA may be any one or more of 3-methyl-uridine (m^3U), 5-methoxy-uridine (mo^5U), 5-aza-uridine, 6-aza-uridine, 2-thio-5-aza-uridine, 2-thio-uridine (s^2U), 4-thio-uridine (s^4U), 4-thio-pseudouridine, 2-thio-pseudouridine, 5-hydroxy-uridine (ho^5U), 5-aminoallyl-uridine, 5-halo-uridine (e.g., 5-iodo-uridine or 5-bromo-uridine), uridine 5-oxyacetic acid (cmo^5U), uridine 5-oxyacetic acid methyl ester ($mcmo^5U$), 5-carboxymethyl-uridine (cm^5U), 1-carboxymethyl-pseudouridine, 5-carboxyhydroxymethyl-uridine (chm^5U), 5-carboxyhydroxymethyl-uridine methyl ester ($mchm^5U$), 5-methoxycarbonylmethyl-uridine (mcm^5U), 5-methoxycarbonylmethyl-2-thio-uridine (mcm^5s^2U), 5-aminomethyl-2-thio-uridine (nm^5s^2U), 5-methylaminomethyl-uridine (mnm^5U), 1-ethyl-pseudouridine, 5-methylaminomethyl-2-thio-uridine (mnm^5s^2U), 5-methylaminomethyl-2-seleno-uridine (mnm^5se^2U), 5-carbamoylmethyl-uridine (ncm^5U), 5-carboxymethylaminomethyl-uridine ($cmnm^5U$), 5-carboxymethylaminomethyl-2-thio-uridine ($cmnm^5s^2U$), 5-propynyl-uridine, 1-propynyl-pseudouridine, 5-taurinomethyl-uridine (τm^5U), 1-taurinomethyl-pseudouridine, 5-taurinomethyl-2-thio-uridine($\tau m5s2U$), 1-taurinomethyl-4-thio-pseudouridine), 5-methyl-2-thio-uridine (m^5s^2U), 1-methyl-4-thio-pseudouridine ($m^1s^4\psi$), 4-thio-1-methyl-pseudouridine, 3-methyl-pseudouridine ($m^3\psi$), 2-thio-1-methyl-pseudouridine, 1-methyl-1-deaza-pseudouridine, 2-thio-1-methyl-1-deaza-pseudouridine, dihydrouridine (D), dihydropseudouridine, 5,6-dihydrouridine, 5-methyl-dihydrouridine (m^5D), 2-thio-dihydrouridine, 2-thio-dihydropseudouridine, 2-methoxy-uridine, 2-methoxy-4-thio-uridine, 4-methoxy-pseudouridine, 4-methoxy-2-thio-pseudouridine, N1-methyl-pseudouridine, 3-(3-amino-3-carboxypropyl)uridine (acp^3U), 1-methyl-3-(3-amino-3-carboxypropyl)pseudouridine ($acp^3\ \psi$), 5-(isopentenylaminomethyl)uridine (inm^5U), 5-(isopentenylaminomethyl)-2-thio-uridine (inm^5s^2U), α-thio-uridine, 2'-O-methyl-uridine (Um), 5,2'-O-dimethyl-uridine (m^5Um), 2'-O-methyl-pseudouridine (ψm), 2-thio-2'-O-methyl-uridine (s^2Um), 5-methoxycarbonylmethyl-2'-O-methyl-uridine (mcm^5Um), 5-carbamoylmethyl-2'-O-methyl-uridine (ncm^5Um), 5-carboxymethylaminomethyl-2'-O-methyl-uridine ($cmnm^5Um$), 3,2'-O-dimethyl-uridine (m^3Um), 5-(isopentenylaminomethyl)-2'-O-methyl-uridine (inm^5Um), 1-thio-uridine, deoxythymidine, 2'-F-ara-uridine, 2'-F-uridine, 2'-

CA 03215771 2023-10-17

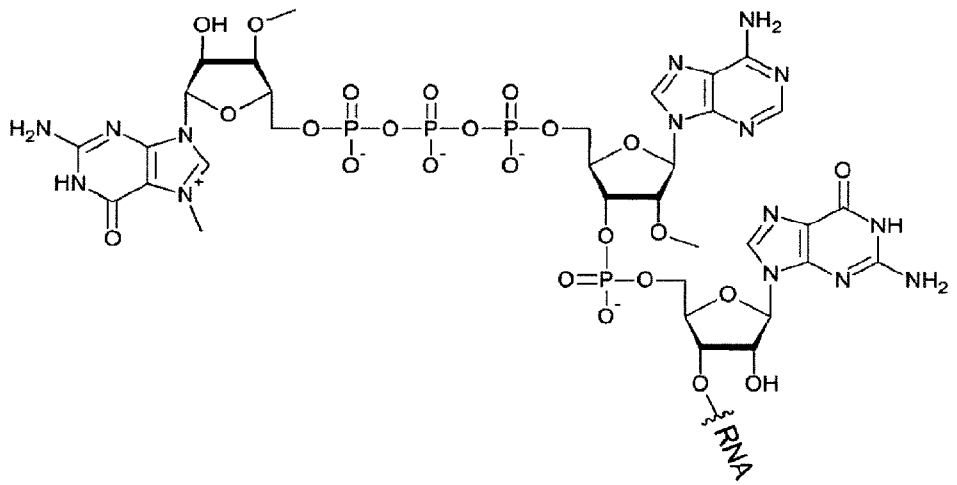
OH-ara-uridine, 5-(2-carbomethoxyvinyl) uridine, 5-[3-(1-E-propenylamino)uridine, or any other modified uridine known in the art.

In one embodiment, the RNA comprises other modified nucleosides or comprises further modified nucleosides, e.g., modified cytidine. For example, in one embodiment, in the RNA 5-methylcytidine is substituted partially or completely, preferably completely, for cytidine. In one embodiment, the RNA comprises 5-methylcytidine and one or more selected from pseudouridine (ψ), N1-methyl-pseudouridine (m1ψ), and 5-methyl-uridine (m5U). In one embodiment, the RNA comprises 5-methylcytidine and N1-methyl-pseudouridine (m1ψ). In some embodiments, the RNA comprises 5-methylcytidine in place of each cytidine and N1-methyl-pseudouridine (m1ψ) in place of each uridine.

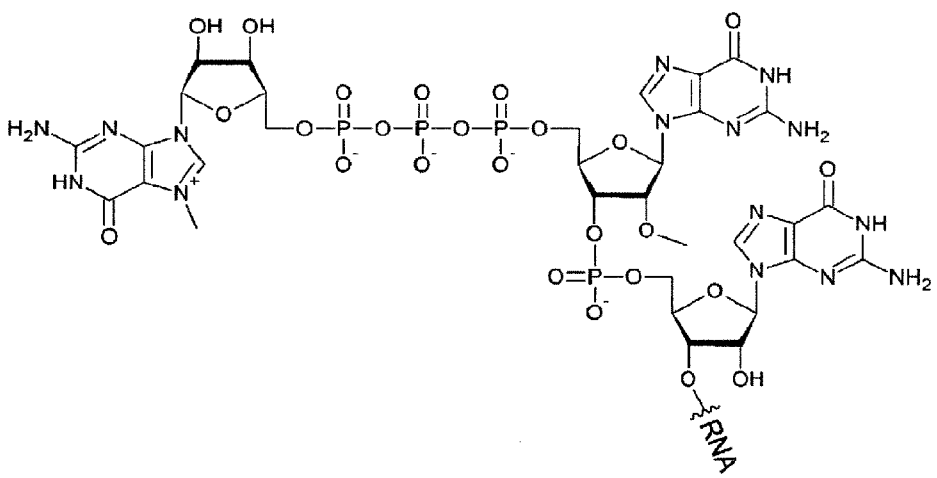
In some embodiments, the RNA according to the present disclosure comprises a 5'-cap. In one embodiment, the RNA of the present disclosure does not have uncapped 5'-triphosphates. In one embodiment, the RNA may be modified by a 5'- cap analog. The term "5'-cap" refers to a structure found on the 5'-end of an mRNA molecule and generally consists of a guanosine nucleotide connected to the mRNA via a 5'- to 5'-triphosphate linkage. In one embodiment, this guanosine is methylated at the 7-position. Providing an RNA with a 5'-cap or 5'-cap analog may be achieved by *in vitro* transcription, in which the 5'-cap is co-transcriptionally expressed into the RNA strand, or may be attached to RNA post-transcriptionally using capping enzymes. In some embodiments, the building block cap for RNA is $m_2^{7,3'\text{-}O}Gppp(m_1^{2'\text{-}O})ApG$ (also sometimes referred to as $m_2^{7,3'O}G(5')ppp(5')m^{2'\text{-}O}ApG$), which has the following structure:



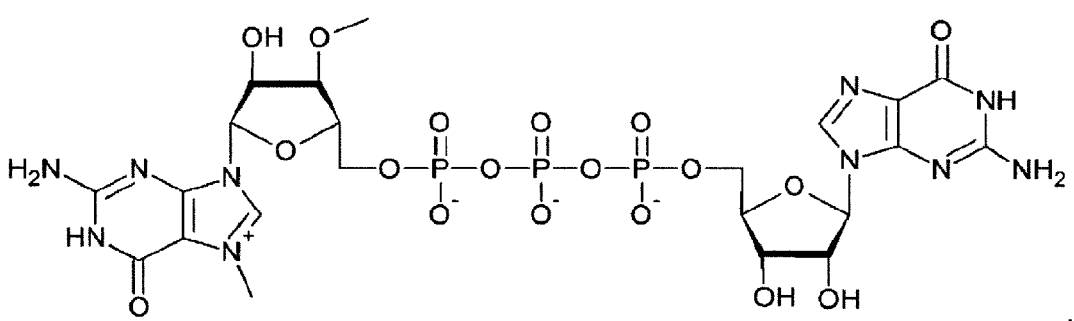
Below is an exemplary Cap1 RNA, which comprises RNA and $m_2^{7,3'O}G(5')ppp(5')m^{2'\text{-}O}ApG$:



Below is another exemplary Cap1 RNA (no cap analog):

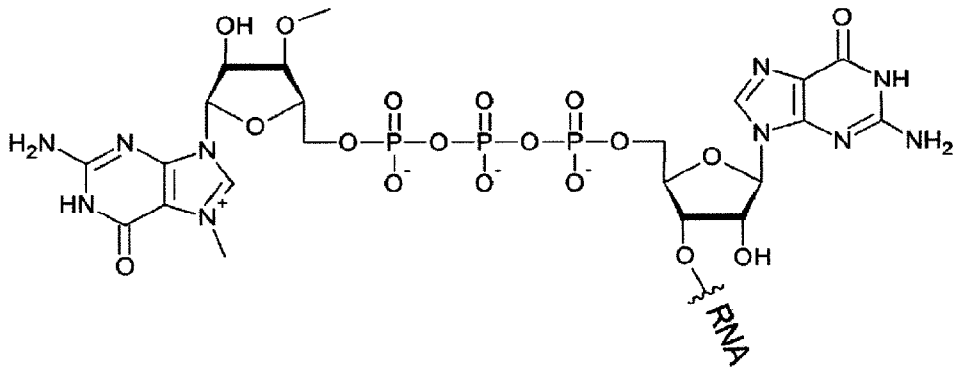


In some embodiments, the RNA is modified with "Cap0" structures using, in one embodiment, the cap analog anti-reverse cap (ARCA Cap ($m_2^{7,3'O}$G(5')ppp(5')G)) with the structure:

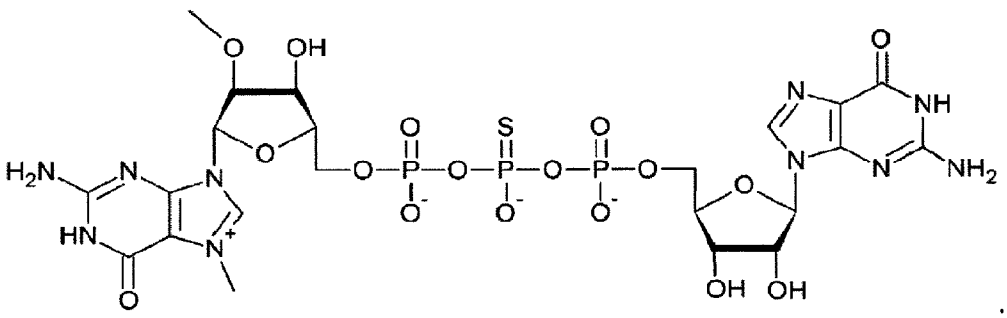


Below is an exemplary Cap0 RNA comprising RNA and $m_2^{7,3'O}$G(5')ppp(5')G:

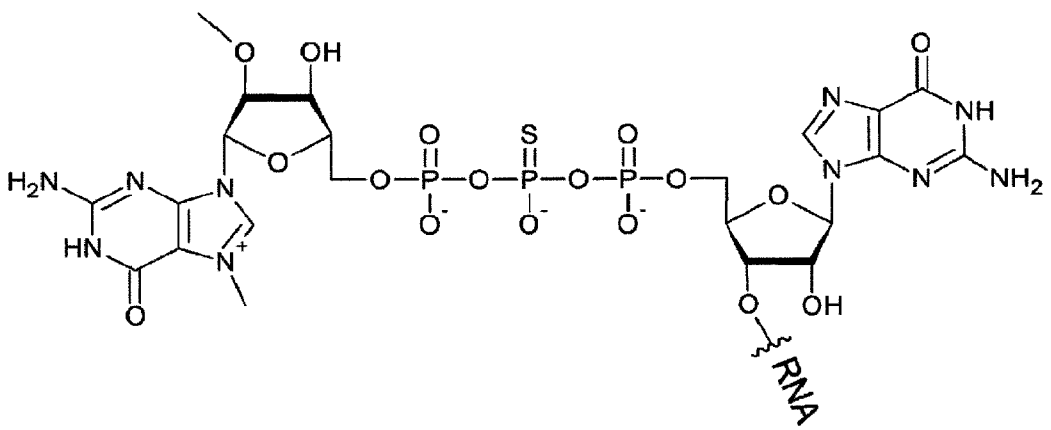
CA 03215771 2023-10-17



In some embodiments, the "Cap0" structures are generated using the cap analog Beta-S-ARCA ($m_2^{7,2`O}$G(5')ppSp(5')G) with the structure:



Below is an exemplary Cap0 RNA comprising Beta-S-ARCA ($m_2^{7,2`O}$G(5')ppSp(5')G) and RNA:



The "D1" diastereomer of beta-S-ARCA or "beta-S-ARCA(D1)" is the diastereomer of beta-S-ARCA which elutes first on an HPLC column compared to the D2 diastereomer of beta-S-ARCA (beta-S-ARCA(D2)) and thus exhibits a shorter retention time (cf., WO 2011/015347, herein incorporated by reference).

A particularly preferred cap is beta-S-ARCA(D1) ($m_2^{7,2'-O}$GppSpG) or $m_2^{7,3'-O}$Gppp($m_1^{2'-O}$)ApG.

In some embodiments, RNA according to the present disclosure comprises a 5'-UTR and/or a 3'-UTR. The term "untranslated region" or "UTR" relates to a region in a DNA molecule which is transcribed but is not translated into an amino acid sequence, or to the corresponding region in an RNA molecule, such as an mRNA molecule. An untranslated region (UTR) can be present 5' (upstream) of an open reading frame (5'-UTR) and/or 3' (downstream) of an open reading frame (3'-UTR). A 5'-UTR, if present, is located at the 5' end, upstream of the start codon of a protein-encoding region. A 5'-UTR is downstream of the 5'-cap (if present), e.g. directly adjacent to the 5'-cap. A 3'-UTR, if present, is located at the 3' end, downstream of the termination codon of a protein-encoding region, but the term "3'-UTR" does preferably not include the poly(A) sequence. Thus, the 3'-UTR is upstream of the poly(A) sequence (if present), e.g. directly adjacent to the poly(A) sequence.

In some embodiments, RNA comprises a 5'-UTR comprising the nucleotide sequence of SEQ ID NO: 12, or a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 12.

In some embodiments, RNA comprises a 3'-UTR comprising the nucleotide sequence of SEQ ID NO: 13, or a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 13.

A particularly preferred 5'-UTR comprises the nucleotide sequence of SEQ ID NO: 12. A particularly preferred 3'-UTR comprises the nucleotide sequence of SEQ ID NO: 13.

In some embodiments, the RNA according to the present disclosure comprises a 3'-poly(A) sequence.

As used herein, the term "poly(A) sequence" or "poly-A tail" refers to an uninterrupted or interrupted sequence of adenylate residues which is typically located at the 3'-end of an RNA molecule. Poly(A) sequences are known to those of skill in the art and may follow the 3'-UTR in the RNAs described herein. An uninterrupted poly(A) sequence is characterized by consecutive adenylate residues. In nature, an uninterrupted poly(A) sequence is typical. RNAs disclosed herein can have a poly(A) sequence attached to the free 3'-end of the RNA by a template-independent RNA polymerase after transcription or a poly(A) sequence encoded by DNA and transcribed by a template-dependent RNA polymerase.

CA 03215771 2023-10-17

It has been demonstrated that a poly(A) sequence of about 120 A nucleotides has a beneficial influence on the levels of RNA in transfected eukaryotic cells, as well as on the levels of protein that is translated from an open reading frame that is present upstream (5') of the poly(A) sequence (*Holtkamp et al.*, 2006, Blood, vol. 108, pp. 4009-4017).

The poly(A) sequence may be of any length. In some embodiments, a poly(A) sequence comprises, essentially consists of, or consists of at least 20, at least 30, at least 40, at least 80, or at least 100 and up to 500, up to 400, up to 300, up to 200, or up to 150 A nucleotides, and, in particular, about 120 A nucleotides. In this context, "essentially consists of" means that most nucleotides in the poly(A) sequence, typically at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% by number of nucleotides in the poly(A) sequence are A nucleotides, but permits that remaining nucleotides are nucleotides other than A nucleotides, such as U nucleotides (uridylate), G nucleotides (guanylate), or C nucleotides (cytidylate). In this context, "consists of" means that all nucleotides in the poly(A) sequence, i.e., 100% by number of nucleotides in the poly(A) sequence, are A nucleotides. The term "A nucleotide" or "A" refers to adenylate.

In some embodiments, a poly(A) sequence is attached during RNA transcription, e.g., during preparation of *in vitro* transcribed RNA, based on a DNA template comprising repeated dT nucleotides (deoxythymidylate) in the strand complementary to the coding strand. The DNA sequence encoding a poly(A) sequence (coding strand) is referred to as poly(A) cassette.

In some embodiments, the poly(A) cassette present in the coding strand of DNA essentially consists of dA nucleotides, but is interrupted by a random sequence of the four nucleotides (dA, dC, dG, and dT). Such random sequence may be 5 to 50, 10 to 30, or 10 to 20 nucleotides in length. Such a cassette is disclosed in WO 2016/005324 A1, hereby incorporated by reference. Any poly(A) cassette disclosed in WO 2016/005324 A1 may be used in the present invention. A poly(A) cassette that essentially consists of dA nucleotides, but is interrupted by a random sequence having an equal distribution of the four nucleotides (dA, dC, dG, dT) and having a length of e.g., 5 to 50 nucleotides shows, on DNA level, constant propagation of plasmid DNA in *E. coli* and is still associated, on RNA level, with the beneficial properties with respect to supporting RNA stability and translational efficiency is encompassed. Consequently, in some embodiments, the poly(A) sequence contained in an RNA molecule described herein

CA 03215771 2023-10-17

essentially consists of A nucleotides, but is interrupted by a random sequence of the four nucleotides (A, C, G, U). Such random sequence may be 5 to 50, 10 to 30, or 10 to 20 nucleotides in length.

In some embodiments, no nucleotides other than A nucleotides flank a poly(A) sequence at its 3'-end, i.e., the poly(A) sequence is not masked or followed at its 3'-end by a nucleotide other than A.

In some embodiments, the poly(A) sequence may comprise at least 20, at least 30, at least 40, at least 80, or at least 100 and up to 500, up to 400, up to 300, up to 200, or up to 150 nucleotides. In some embodiments, the poly(A) sequence may essentially consist of at least 20, at least 30, at least 40, at least 80, or at least 100 and up to 500, up to 400, up to 300, up to 200, or up to 150 nucleotides. In some embodiments, the poly(A) sequence may consist of at least 20, at least 30, at least 40, at least 80, or at least 100 and up to 500, up to 400, up to 300, up to 200, or up to 150 nucleotides. In some embodiments, the poly(A) sequence comprises at least 100 nucleotides. In some embodiments, the poly(A) sequence comprises about 150 nucleotides. In some embodiments, the poly(A) sequence comprises about 120 nucleotides.

In some embodiments, RNA comprises a poly(A) sequence comprising the nucleotide sequence of SEQ ID NO: 14, or a nucleotide sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the nucleotide sequence of SEQ ID NO: 14.

A particularly preferred poly(A) sequence comprises comprises the nucleotide sequence of SEQ ID NO: 14.

According to the disclosure, vaccine antigen is preferably administered as single-stranded, 5'-capped mRNA that is translated into the respective protein upon entering cells of a subject being administered the RNA. Preferably, the RNA contains structural elements optimized for maximal efficacy of the RNA with respect to stability and translational efficiency (5'-cap, 5'-UTR, 3'-UTR, poly(A) sequence).

In one embodiment, beta-S-ARCA(D1) is utilized as specific capping structure at the 5'-end of the RNA. In one embodiment, $m_2^{7,3'-O}Gppp(m_1^{2'-O})ApG$ is utilized as specific capping structure at the 5'-end of the RNA. In one embodiment, the 5'-UTR sequence is derived from the human alpha-globin mRNA and optionally has an optimized 'Kozak sequence' to increase translational

CA 03215771 2023-10-17

efficiency. In one embodiment, a combination of two sequence elements (FI element) derived from the "amino terminal enhancer of split" (AES) mRNA (called F) and the mitochondrial encoded 12S ribosomal RNA (called I) are placed between the coding sequence and the poly(A) sequence to assure higher maximum protein levels and prolonged persistence of the mRNA. In one embodiment, two re-iterated 3'-UTRs derived from the human beta-globin mRNA are placed between the coding sequence and the poly(A) sequence to assure higher maximum protein levels and prolonged persistence of the mRNA. In one embodiment, a poly(A) sequence measuring 110 nucleotides in length, consisting of a stretch of 30 adenosine residues, followed by a 10 nucleotide linker sequence and another 70 adenosine residues is used. This poly(A) sequence was designed to enhance RNA stability and translational efficiency.

In one embodiment of all aspects of the invention, RNA encoding a vaccine antigen is expressed in cells of the subject treated to provide the vaccine antigen. In one embodiment of all aspects of the invention, the RNA is transiently expressed in cells of the subject. In one embodiment of all aspects of the invention, the RNA is *in vitro* transcribed RNA. In one embodiment of all aspects of the invention, expression of the vaccine antigen is at the cell surface. In one embodiment of all aspects of the invention, the vaccine antigen is expressed and presented in the context of MHC. In one embodiment of all aspects of the invention, expression of the vaccine antigen is into the extracellular space, i.e., the vaccine antigen is secreted.

In the context of the present disclosure, the term "transcription" relates to a process, wherein the genetic code in a DNA sequence is transcribed into RNA. Subsequently, the RNA may be translated into peptide or protein.

According to the present invention, the term "transcription" comprises "*in vitro* transcription", wherein the term "*in vitro* transcription" relates to a process wherein RNA, in particular mRNA, is *in vitro* synthesized in a cell-free system, preferably using appropriate cell extracts. Preferably, cloning vectors are applied for the generation of transcripts. These cloning vectors are generally designated as transcription vectors and are according to the present invention encompassed by the term "vector". According to the present invention, the RNA used in the present invention preferably is *in vitro* transcribed RNA (IVT-RNA) and may be obtained by *in*

CA 03215771 2023-10-17

vitro transcription of an appropriate DNA template. The promoter for controlling transcription can be any promoter for any RNA polymerase. Particular examples of RNA polymerases are the T7, T3, and SP6 RNA polymerases. Preferably, the *in vitro* transcription according to the invention is controlled by a T7 or SP6 promoter. A DNA template for *in vitro* transcription may be obtained by cloning of a nucleic acid, in particular cDNA, and introducing it into an appropriate vector for *in vitro* transcription. The cDNA may be obtained by reverse transcription of RNA.

With respect to RNA, the term "expression" or "translation" relates to the process in the ribosomes of a cell by which a strand of mRNA directs the assembly of a sequence of amino acids to make a peptide or protein.

In one embodiment, after administration of the RNA described herein, e.g., formulated as RNA lipid particles, at least a portion of the RNA is delivered to a target cell. In one embodiment, at least a portion of the RNA is delivered to the cytosol of the target cell. In one embodiment, the RNA is translated by the target cell to produce the peptide or protein it enodes. In one embodiment, the target cell is a spleen cell. In one embodiment, the target cell is an antigen presenting cell such as a professional antigen presenting cell in the spleen. In one embodiment, the target cell is a dendritic cell or macrophage. RNA particles such as RNA lipid particles described herein may be used for delivering RNA to such target cell. Accordingly, the present disclosure also relates to a method for delivering RNA to a target cell in a subject comprising the administration of the RNA particles described herein to the subject. In one embodiment, the RNA is delivered to the cytosol of the target cell. In one embodiment, the RNA is translated by the target cell to produce the peptide or protein encoded by the RNA.

"Encoding" refers to the inherent property of specific sequences of nucleotides in a polynucleotide, such as a gene, a cDNA, or an mRNA, to serve as templates for synthesis of other polymers and macromolecules in biological processes having either a defined sequence of nucleotides (i.e., rRNA, tRNA and mRNA) or a defined sequence of amino acids and the biological properties resulting therefrom. Thus, a gene encodes a protein if transcription and translation of mRNA corresponding to that gene produces the protein in a cell or other biological system. Both the coding strand, the nucleotide sequence of which is identical to the mRNA sequence and is usually provided in sequence listings, and the non-coding strand, used

CA 03215771 2023-10-17

as the template for transcription of a gene or cDNA, can be referred to as encoding the protein or other product of that gene or cDNA.

In one embodiment, nucleic acid compositions described herein, e.g., compositions comprising a lipid nanoparticle encapsulated mRNA are characterized by (e.g., when administered to a subject) sustained expression of an encoded polypeptide. For example, in some embodiments, such compositions are characterized in that, when administered to a human, they achieve detectable polypeptide expression in a biological sample (e.g., serum) from such human and, in some embodiments, such expression persists for a period of time that is at least at least 36 hours or longer, including, e.g., at least 48 hours, at least 60 hours, at least 72 hours, at least 96 hours, at least 120 hours, at least 148 hours, or longer.

In one embodiment, the RNA encoding vaccine antigen to be administered according to the invention is non-immunogenic. RNA encoding immunostimulant may be administered according to the invention to provide an adjuvant effect. The RNA encoding immunostimulant may be standard RNA or non-immunogenic RNA.

The term "non-immunogenic RNA" as used herein refers to RNA that does not induce a response by the immune system upon administration, e.g., to a mammal, or induces a weaker response than would have been induced by the same RNA that differs only in that it has not been subjected to the modifications and treatments that render the non-immunogenic RNA non-immunogenic, i.e., than would have been induced by standard RNA (stdRNA). In one preferred embodiment, non-immunogenic RNA, which is also termed modified RNA (modRNA) herein, is rendered non-immunogenic by incorporating modified nucleosides suppressing RNA-mediated activation of innate immune receptors into the RNA and removing double-stranded RNA (dsRNA).

For rendering the non-immunogenic RNA non-immunogenic by the incorporation of modified nucleosides, any modified nucleoside may be used as long as it lowers or suppresses immunogenicity of the RNA. Particularly preferred are modified nucleosides that suppress RNA-mediated activation of innate immune receptors. In one embodiment, the modified nucleosides comprises a replacement of one or more uridines with a nucleoside comprising a modified nucleobase. In one embodiment, the modified nucleobase is a modified uracil. In one embodiment, the nucleoside comprising a modified nucleobase is selected from the group

consisting of 3-methyl-uridine (m^3U), 5-methoxy-uridine (mo^5U), 5-aza-uridine, 6-aza-uridine, 2-thio-5-aza-uridine, 2-thio-uridine (s^2U), 4-thio-uridine (s^4U), 4-thio-pseudouridine, 2-thio-pseudouridine, 5-hydroxy-uridine (ho^5U), 5-aminoallyl-uridine, 5-halo-uridine (e.g., 5-iodo-uridine or 5-bromo-uridine), uridine 5-oxyacetic acid (cmo^5U), uridine 5-oxyacetic acid methyl ester ($mcmo^5U$), 5-carboxymethyl-uridine (cm^5U), 1-carboxymethyl-pseudouridine, 5-carboxyhydroxymethyl-uridine (chm^5U), 5-carboxyhydroxymethyl-uridine methyl ester ($mchm^5U$), 5-methoxycarbonylmethyl-uridine (mcm^5U), 5-methoxycarbonylmethyl-2-thio-uridine (mcm^5s^2U), 5-aminomethyl-2-thio-uridine (nm^5s^2U), 5-methylaminomethyl-uridine (mnm^5U), 1-ethyl-pseudouridine, 5-methylaminomethyl-2-thio-uridine (mnm^5s^2U), 5-methylaminomethyl-2-seleno-uridine (mnm^5se^2U), 5-carbamoylmethyl-uridine (ncm^5U), 5-carboxymethylaminomethyl-uridine ($cmnm^5U$), 5-carboxymethylaminomethyl-2-thio-uridine ($cmnm^5s^2U$), 5-propynyl-uridine, 1-propynyl-pseudouridine, 5-taurinomethyl-uridine (τm^5U), 1-taurinomethyl-pseudouridine, 5-taurinomethyl-2-thio-uridine($\tau m5s2U$), 1-taurinomethyl-4-thio-pseudouridine), 5-methyl-2-thio-uridine (m^5s^2U), 1-methyl-4-thio-pseudouridine ($m^1s^4\psi$), 4-thio-1-methyl-pseudouridine, 3-methyl-pseudouridine ($m^3\psi$), 2-thio-1-methyl-pseudouridine, 1-methyl-1-deaza-pseudouridine, 2-thio-1-methyl-1-deaza-pseudouridine, dihydrouridine (D), dihydropseudouridine, 5,6-dihydrouridine, 5-methyl-dihydrouridine (m^5D), 2-thio-dihydrouridine, 2-thio-dihydropseudouridine, 2-methoxy-uridine, 2-methoxy-4-thio-uridine, 4-methoxy-pseudouridine, 4-methoxy-2-thio-pseudouridine, N1-methyl-pseudouridine, 3-(3-amino-3-carboxypropyl)uridine (acp^3U), 1-methyl-3-(3-amino-3-carboxypropyl)pseudouridine ($acp^3\ \psi$), 5-(isopentenylaminomethyl)uridine (inm^5U), 5-(isopentenylaminomethyl)-2-thio-uridine (inm^5s^2U), α-thio-uridine, 2'-O-methyl-uridine (Um), 5,2'-O-dimethyl-uridine (m^5Um), 2'-O-methyl-pseudouridine (ψm), 2-thio-2'-O-methyl-uridine (s^2Um), 5-methoxycarbonylmethyl-2'-O-methyl-uridine (mcm^5Um), 5-carbamoylmethyl-2'-O-methyl-uridine (ncm^5Um), 5-carboxymethylaminomethyl-2'-O-methyl-uridine ($cmnm^5Um$), 3,2'-O-dimethyl-uridine (m^3Um), 5-(isopentenylaminomethyl)-2'-O-methyl-uridine (inm^5Um), 1-thio-uridine, deoxythymidine, 2'-F-ara-uridine, 2'-F-uridine, 2'-OH-ara-uridine, 5-(2-carbomethoxyvinyl) uridine, and 5-[3-(1-E-propenylamino)uridine. In one particularly preferred embodiment, the nucleoside comprising a modified nucleobase is

CA 03215771 2023-10-17

pseudouridine (ψ), N1-methyl-pseudouridine (m1ψ) or 5-methyl-uridine (m5U), in particular N1-methyl-pseudouridine.

In one embodiment, the replacement of one or more uridines with a nucleoside comprising a modified nucleobase comprises a replacement of at least 1%, at least 2%, at least 3%, at least 4%, at least 5%, at least 10%, at least 25%, at least 50%, at least 75%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or 100% of the uridines.

During synthesis of mRNA by *in vitro* transcription (IVT) using T7 RNA polymerase significant amounts of aberrant products, including double-stranded RNA (dsRNA) are produced due to unconventional activity of the enzyme. dsRNA induces inflammatory cytokines and activates effector enzymes leading to protein synthesis inhibition. dsRNA can be removed from RNA such as IVT RNA, for example, by ion-pair reversed phase HPLC using a non-porous or porous C-18 polystyrene-divinylbenzene (PS-DVB) matrix. Alternatively, an enzymatic based method using E. coli RNaseIII that specifically hydrolyzes dsRNA but not ssRNA, thereby eliminating dsRNA contaminants from IVT RNA preparations can be used. Furthermore, dsRNA can be separated from ssRNA by using a cellulose material. In one embodiment, an RNA preparation is contacted with a cellulose material and the ssRNA is separated from the cellulose material under conditions which allow binding of dsRNA to the cellulose material and do not allow binding of ssRNA to the cellulose material.

As the term is used herein, "remove" or "removal" refers to the characteristic of a population of first substances, such as non-immunogenic RNA, being separated from the proximity of a population of second substances, such as dsRNA, wherein the population of first substances is not necessarily devoid of the second substance, and the population of second substances is not necessarily devoid of the first substance. However, a population of first substances characterized by the removal of a population of second substances has a measurably lower content of second substances as compared to the non-separated mixture of first and second substances.

In one embodiment, the removal of dsRNA from non-immunogenic RNA comprises a removal of dsRNA such that less than 10%, less than 5%, less than 4%, less than 3%, less than 2%, less than 1%, less than 0.5%, less than 0.3%, or less than 0.1% of the RNA in the non-immunogenic RNA composition is dsRNA. In one embodiment, the non-immunogenic RNA is free or

CA 03215771 2023-10-17

essentially free of dsRNA. In some embodiments, the non-immunogenic RNA composition comprises a purified preparation of single-stranded nucleoside modified RNA. For example, in some embodiments, the purified preparation of single-stranded nucleoside modified RNA is substantially free of double stranded RNA (dsRNA). In some embodiments, the purified preparation is at least 90%, at least 91%, at least 92%, at least 93 %, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5%, or at least 99.9% single stranded nucleoside modified RNA, relative to all other nucleic acid molecules (DNA, dsRNA, etc.).

In one embodiment, the non-immunogenic RNA is translated in a cell more efficiently than standard RNA with the same sequence. In one embodiment, translation is enhanced by a factor of 2-fold relative to its unmodified counterpart. In one embodiment, translation is enhanced by a 3-fold factor. In one embodiment, translation is enhanced by a 4-fold factor. In one embodiment, translation is enhanced by a 5-fold factor. In one embodiment, translation is enhanced by a 6-fold factor. In one embodiment, translation is enhanced by a 7-fold factor. In one embodiment, translation is enhanced by an 8-fold factor. In one embodiment, translation is enhanced by a 9-fold factor. In one embodiment, translation is enhanced by a 10-fold factor. In one embodiment, translation is enhanced by a 15-fold factor. In one embodiment, translation is enhanced by a 20-fold factor. In one embodiment, translation is enhanced by a 50-fold factor. In one embodiment, translation is enhanced by a 100-fold factor. In one embodiment, translation is enhanced by a 200-fold factor. In one embodiment, translation is enhanced by a 500-fold factor. In one embodiment, translation is enhanced by a 1000-fold factor. In one embodiment, translation is enhanced by a 2000-fold factor. In one embodiment, the factor is 10-1000-fold. In one embodiment, the factor is 10-100-fold. In one embodiment, the factor is 10-200-fold. In one embodiment, the factor is 10-300-fold. In one embodiment, the factor is 10-500-fold. In one embodiment, the factor is 20-1000-fold. In one embodiment, the factor is 30-1000-fold. In one embodiment, the factor is 50-1000-fold. In one embodiment, the factor is 100-1000-fold. In one embodiment, the factor is 200-1000-fold. In one embodiment, translation is enhanced by any other significant amount or range of amounts.

In one embodiment, the non-immunogenic RNA exhibits significantly less innate immunogenicity than standard RNA with the same sequence. In one embodiment, the non-immunogenic RNA exhibits an innate immune response that is 2-fold less than its unmodified counterpart. In one embodiment, innate immunogenicity is reduced by a 3-fold factor. In one embodiment, innate immunogenicity is reduced by a 4-fold factor. In one embodiment, innate immunogenicity is reduced by a 5-fold factor. In one embodiment, innate immunogenicity is reduced by a 6-fold factor. In one embodiment, innate immunogenicity is reduced by a 7-fold factor. In one embodiment, innate immunogenicity is reduced by a 8-fold factor. In one embodiment, innate immunogenicity is reduced by a 9-fold factor. In one embodiment, innate immunogenicity is reduced by a 10-fold factor. In one embodiment, innate immunogenicity is reduced by a 15-fold factor. In one embodiment, innate immunogenicity is reduced by a 20-fold factor. In one embodiment, innate immunogenicity is reduced by a 50-fold factor. In one embodiment, innate immunogenicity is reduced by a 100-fold factor. In one embodiment, innate immunogenicity is reduced by a 200-fold factor. In one embodiment, innate immunogenicity is reduced by a 500-fold factor. In one embodiment, innate immunogenicity is reduced by a 1000-fold factor. In one embodiment, innate immunogenicity is reduced by a 2000-fold factor.

The term "exhibits significantly less innate immunogenicity" refers to a detectable decrease in innate immunogenicity. In one embodiment, the term refers to a decrease such that an effective amount of the non-immunogenic RNA can be administered without triggering a detectable innate immune response. In one embodiment, the term refers to a decrease such that the non-immunogenic RNA can be repeatedly administered without eliciting an innate immune response sufficient to detectably reduce production of the protein encoded by the non-immunogenic RNA. In one embodiment, the decrease is such that the non-immunogenic RNA can be repeatedly administered without eliciting an innate immune response sufficient to eliminate detectable production of the protein encoded by the non-immunogenic RNA.

"Immunogenicity" is the ability of a foreign substance, such as RNA, to provoke an immune response in the body of a human or other animal. The innate immune system is the component of the immune system that is relatively unspecific and immediate. It is one of two main components of the vertebrate immune system, along with the adaptive immune system.

As used herein "endogenous" refers to any material from or produced inside an organism, cell, tissue or system.

As used herein, the term "exogenous" refers to any material introduced from or produced outside an organism, cell, tissue or system.

The term "expression" as used herein is defined as the transcription and/or translation of a particular nucleotide sequence.

As used herein, the terms "linked," "fused", or "fusion" are used interchangeably. These terms refer to the joining together of two or more elements or components or domains.

Codon-optimization / Increase in G/C content

In some embodiment, the amino acid sequence comprising a polyspecific virus protein amino acid sequence described herein is encoded by a coding sequence which is codon-optimized and/or the G/C content of which is increased compared to wild type coding sequence. This also includes embodiments, wherein one or more sequence regions of the coding sequence are codon-optimized and/or increased in the G/C content compared to the corresponding sequence regions of the wild type coding sequence. In one embodiment, the codon-optimization and/or the increase in the G/C content preferably does not change the sequence of the encoded amino acid sequence.

The term "codon-optimized" refers to the alteration of codons in the coding region of a nucleic acid molecule to reflect the typical codon usage of a host organism without preferably altering the amino acid sequence encoded by the nucleic acid molecule. Within the context of the present invention, coding regions are preferably codon-optimized for optimal expression in a subject to be treated using the RNA molecules described herein. Codon-optimization is based on the finding that the translation efficiency is also determined by a different frequency in the occurrence of tRNAs in cells. Thus, the sequence of RNA may be modified such that codons for which frequently occurring tRNAs are available are inserted in place of "rare codons".

In some embodiments of the invention, the guanosine/cytosine (G/C) content of the coding region of the RNA described herein is increased compared to the G/C content of the corresponding coding sequence of the wild type RNA, wherein the amino acid sequence encoded by the RNA is preferably not modified compared to the amino acid sequence

encoded by the wild type RNA. This modification of the RNA sequence is based on the fact that the sequence of any RNA region to be translated is important for efficient translation of that mRNA. Sequences having an increased G (guanosine)/C (cytosine) content are more stable than sequences having an increased A (adenosine)/U (uracil) content. In respect to the fact that several codons code for one and the same amino acid (so-called degeneration of the genetic code), the most favourable codons for the stability can be determined (so-called alternative codon usage). Depending on the amino acid to be encoded by the RNA, there are various possibilities for modification of the RNA sequence, compared to its wild type sequence. In particular, codons which contain A and/or U nucleotides can be modified by substituting these codons by other codons, which code for the same amino acids but contain no A and/or U or contain a lower content of A and/or U nucleotides.

In various embodiments, the G/C content of the coding region of the RNA described herein is increased by at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 55%, or even more compared to the G/C content of the coding region of the wild type RNA.

Embodiments of administered RNAs

In some embodiments, the present disclosure provides an RNA (e.g., mRNA) comprising an open reading frame encoding a polypeptide that comprises at least a portion of a SARS-CoV-2 S protein, which at least a portion of a SARS-CoV-2 S protein is modified to incorporate amino acid modifications found in other SARS-CoV-2 S protein variants. The RNA may comprise a plurality of different RNA molecules encoding polypeptides with different of said modifications. The RNA is suitable for intracellular expression of the polypeptide. In some embodiments, such an encoded polypeptide does not comprise a sequence corresponding to the complete S protein. In some embodiments, the encoded polypeptide comprises a sequence that corresponds to the receptor binding domain (RBD). In some embodiments, such an RNA (e.g., mRNA) may be complexed by a (poly)cationic polymer, polyplex(es), protein(s) or peptide(s). In some embodiments, such an RNA may be formulated in a lipid nanoparticle (e.g., ones described herein). In some embodiments, such an RNA (e.g., mRNA) may be particularly useful and/or effective for use as or in an immunogenic composition (e.g., a vaccine), and/or for achieving immunological effects as described herein (e.g., generation of

CA 03215771 2023-10-17

SARS-CoV-2 neutralizing antibodies, and/or T cell responses (e.g., CD4+ and/or CD8+ T cell responses)). In some embodiments, such an RNA (e.g., mRNA) may be useful for vaccinating humans (including, e.g., humans known to have been exposed and/or infected by SARS-CoV-2, and/or humans not known to have been exposed to SARS-CoV-2).

In one embodiment, mRNA constructs described herein comprise a nucleic acid sequence that encodes an amino acid sequence corresponding to a full-length SARS-CoV-2 Spike protein (e.g., including embodiments in which such encoded amino acid sequence may comprise at least one or more amino acid substitutions, e.g., proline substitutions as described herein, and/or embodiments in which the mRNA sequence is codon-optimized e.g., for mammalian, e.g., human, subjects). In one embodiment, mRNA constructs that encode an amino acid sequence corresponding to a full-length SARS-CoV-2 S protein may be particularly useful and/or effective for use as or in an immunogenic composition (e.g., a vaccine) in particular subject population (e.g., particular age populations, e.g., age populations as described herein). In some embodiments, compositions or medical preparations described herein comprise RNA encoding an amino acid sequence that is a modified variant of an amino acid sequence comprising SARS-CoV-2 S protein, or an immunogenic fragment of the SARS-CoV-2 S protein. Likewise, methods described herein comprise administration of such RNA.

The active platform for use herein may be based on an antigen-coding RNA vaccine to induce robust neutralising antibodies and accompanying/concomitant T cell response to achieve protective immunization with preferably minimal vaccine doses. The RNA administered is preferably *in-vitro* transcribed RNA.

Three different RNA platforms are particularly preferred, namely non-modified uridine containing mRNA (uRNA), nucleoside modified mRNA (modRNA) and self-amplifying RNA (saRNA). In one particularly preferred embodiment, the RNA is in vitro transcribed RNA.

In the following, embodiments of these three different RNA platforms are described, wherein certain terms used when describing elements thereof have the following meanings:

S1S2 protein/S1S2 RBD: Sequences encoding the respective (modified) antigen of SARS-CoV-2.

CA 03215771 2023-10-17

nsP1, nsP2, nsP3, and nsP4: Wildtype sequences encoding the Venezuelan equine encephalitis virus (VEEV) RNA-dependent RNA polymerase replicase and a subgenomic promotor plus conserved sequence elements supporting replication and translation.

virUTR: Viral untranslated region encoding parts of the subgenomic promotor as well as replication and translation supporting sequence elements.

hAg-Kozak: 5'-UTR sequence of the human alpha-globin mRNA with an optimized 'Kozak sequence' to increase translational efficiency.

Sec: Sec corresponds to the intrinsic S1S2 protein secretory signal peptide (sec), which guides translocation of the nascent polypeptide chain into the endoplasmatic reticulum.

Glycine-serine linker (GS): Sequences coding for short linker peptides predominantly consisting of the amino acids glycine (G) and serine (S), as commonly used for fusion proteins.

Fibritin: Partial sequence of T4 fibritin (foldon), used as artificial trimerization domain.

TM: TM sequence corresponds to the transmembrane part of the S1S2 protein.

FI element: The 3'-UTR is a combination of two sequence elements derived from the "amino terminal enhancer of split" (AES) mRNA (called F) and the mitochondrial encoded 12S ribosomal RNA (called I). These were identified by an *ex vivo* selection process for sequences that confer RNA stability and augment total protein expression.

A30L70: A poly(A)-tail measuring 110 nucleotides in length, consisting of a stretch of 30 adenosine residues, followed by a 10 nucleotide linker sequence and another 70 adenosine residues designed to enhance RNA stability and translational efficiency in dendritic cells.

In general, vaccine RNA described herein may comprise, from 5' to 3', one of the following structures:

Cap-5'-UTR-Vaccine Antigen-Encoding Sequence-3'-UTR-Poly(A)

or

Cap-hAg-Kozak-Vaccine Antigen-Encoding Sequence-FI-A30L70.

In general, a vaccine antigen described herein may comprise, from N-terminus to C-terminus, one of the following structures:

Signal Sequence-RBD-Trimerization Domain

CA 03215771 2023-10-17

or

Signal Sequence-RBD-Trimerization Domain-Transmembrane Domain.

RBD and Trimerization Domain may be separated by a linker, in particular a GS linker such as a linker having the amino acid sequence GSPGSGSGS. Trimerization Domain and Transmembrane Domain may be separated by a linker, in particular a GS linker such as a linker having the amino acid sequence GSGSGS.

Signal Sequence may be a signal sequence as described herein. RBD may be a (modified) RBD domain as described herein. Trimerization Domain may be a trimerization domain as described herein. Transmembrane Domain may be a transmembrane domain as described herein.

In one embodiment,

Signal sequence comprises the amino acid sequence of amino acids 1 to 16 or 1 to 19 of SEQ ID NO: 1 or the amino acid sequence of amino acids 1 to 22 of SEQ ID NO: 31, or an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to this amino acid sequence,

RBD comprises an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to the amino acid sequence of amino acids 327 to 528 of SEQ ID NO: 1,

Trimerization Domain comprises the amino acid sequence of amino acids 3 to 29 of SEQ ID NO: 10 or the amino acid sequence of SEQ ID NO: 10, or an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to this amino acid sequence; and

Transmembrane Domain comprises the amino acid sequence of amino acids 1207 to 1254 of SEQ ID NO: 1, or an amino acid sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, or 80% identity to this amino acid sequence.

In one embodiment,

Signal sequence comprises the amino acid sequence of amino acids 1 to 16 or 1 to 19 of SEQ ID NO: 1 or the amino acid sequence of amino acids 1 to 22 of SEQ ID NO: 31,

Trimerization Domain comprises the amino acid sequence of amino acids 3 to 29 of SEQ ID NO: 10 or the amino acid sequence of SEQ ID NO: 10; and

Transmembrane Domain comprises the amino acid sequence of amino acids 1207 to 1254 of SEQ ID NO: 1.

The RNA described herein or RNA encoding the vaccine antigen described herein may be non-modified uridine containing mRNA (uRNA), nucleoside modified mRNA (modRNA) or self-amplifying RNA (saRNA). In one embodiment, the RNA described herein or RNA encoding the vaccine antigen described herein is nucleoside modified mRNA (modRNA).

Non-modified uridine messenger RNA (uRNA)

The active principle of the non-modified messenger RNA (uRNA) drug substance is a single-stranded mRNA that is translated upon entering a cell. In addition to the sequence encoding the vaccine antigen (i.e. open reading frame), each uRNA preferably contains common structural elements optimized for maximal efficacy of the RNA with respect to stability and translational efficiency (5'-cap, 5'-UTR, 3'-UTR, poly(A)-tail). The preferred 5' cap structure is beta-S-ARCA(D1) ($m_2^{7,2'\text{-}O}$GppSpG). The preferred 5'-UTR and 3'-UTR comprise the nucleotide sequence of SEQ ID NO: 12 and the nucleotide sequence of SEQ ID NO: 13, respectively. The preferred poly(A)-tail comprises the sequence of SEQ ID NO: 14.

Different embodiments of this platform are sequences that are derived from and are modified variants of the following:

RBL063.1 (SEQ ID NO: 15; SEQ ID NO: 7)

Structure	beta-S-ARCA(D1)-hAg-Kozak-S1S2-PP-FI-A30L70
Encoded antigen	Viral spike protein (S1S2 protein) of the SARS-CoV-2 (S1S2 full-length protein, sequence variant)

RBL063.2 (SEQ ID NO: 16; SEQ ID NO: 7)

Structure	beta-S-ARCA(D1)-hAg-Kozak-S1S2-PP-FI-A30L70

CA 03215771 2023-10-17

Encoded antigen	Viral spike protein (S1S2 protein) of the SARS-CoV-2 (S1S2 full-length protein, sequence variant)

BNT162a1; RBL063.3 (SEQ ID NO: 17; SEQ ID NO: 5)

Structure	beta-S-ARCA(D1)-hAg-Kozak-RBD-GS-Fibritin-FI-A30L70
Encoded antigen	Viral spike protein (S protein) of the SARS-CoV-2 (partial sequence, Receptor Binding Domain (RBD) of S1S2 protein)

Figure 3 schematizes the general structure of the antigen-encoding RNAs.

Nucleoside modified messenger RNA (modRNA)

The active principle of the nucleoside modified messenger RNA (modRNA) drug substance is as well a single-stranded mRNA that is translated upon entering a cell. In addition to the sequence encoding the vaccine antigen (i.e. open reading frame), each modRNA contains common structural elements optimized for maximal efficacy of the RNA as the uRNA (5'-cap, 5'-UTR, 3'-UTR, poly(A)-tail). Compared to the uRNA, modRNA contains 1-methyl-pseudouridine instead of uridine. The preferred 5' cap structure is $m_2^{7,3'-O}Gppp(m_1^{2'-O})ApG$. The preferred 5'-UTR and 3'-UTR comprise the nucleotide sequence of SEQ ID NO: 12 and the nucleotide sequence of SEQ ID NO: 13, respectively. The preferred poly(A)-tail comprises the sequence of SEQ ID NO: 14. An additional purification step is applied for modRNA to reduce dsRNA contaminants generated during the in vitro transcription reaction.

Different embodiments of this platform are sequences that are derived from and are modified variants of the following:

BNT162b2; RBP020.1 (SEQ ID NO: 19; SEQ ID NO: 7)

Structure	$m_2^{7,3'-O}Gppp(m_1^{2'-O})ApG$)-hAg-Kozak-S1S2-PP-FI-A30L70
Encoded antigen	Viral spike protein (S1S2 protein) of the SARS-CoV-2 (S1S2 full-length protein, sequence variant)

CA 03215771 2023-10-17

BNT162b2; RBP020.2 (SEQ ID NO: 20; SEQ ID NO: 7)

Structure $m_2^{7,3'-O}$Gppp($m_1^{2'-O}$)ApG)-hAg-Kozak-S1S2-PP-FI-A30L70

Encoded antigen Viral spike protein (S1S2 protein) of the SARS-CoV-2 (S1S2 full-length protein, sequence variant)

BNT162b1; RBP020.3 (SEQ ID NO: 21; SEQ ID NO: 5)

Structure $m_2^{7,3'-O}$Gppp($m_1^{2'-O}$)ApG)-hAg-Kozak-RBD-GS-Fibritin-FI-A30L70

Encoded antigen Viral spike protein (S1S2 protein) of the SARS-CoV-2 (partial sequence, Receptor Binding Domain (RBD) of S1S2 protein fused to fibritin)

Figure 4 schematizes the general structure of the antigen-encoding RNAs.

BNT162b3c (SEQ ID NO: 29; SEQ ID NO: 30)

Structure $m_2^{7,3'-O}$Gppp($m_1^{2'-O}$)ApG-hAg-Kozak-RBD-GS-Fibritin-GS-TM-FI-A30L70

Encoded antigen Viral spike protein (S1S2 protein) of the SARS-CoV-2 (partial sequence, Receptor Binding Domain (RBD) of S1S2 protein fused to Fibritin fused to Transmembrane Domain (TM) of S1S2 protein); intrinsic S1S2 protein secretory signal peptide (aa 1-19) at the N-terminus of the antigen sequence.

BNT162b3d (SEQ ID NO: 31; SEQ ID NO: 32)

Structure $m_2^{7,3'-O}$Gppp($m_1^{2'-O}$)ApG-hAg-Kozak-RBD-GS-Fibritin-GS-TM-FI-A30L70

Encoded antigen Viral spike protein (S1S2 protein) of the SARS-CoV-2 (partial sequence, Receptor Binding Domain (RBD) of S1S2 protein fused to Fibritin fused to Transmembrane Domain (TM) of S1S2 protein); immunoglobulin secretory signal peptide (aa 1-22) at the N-terminus of the antigen sequence.

Self-amplifying RNA (saRNA)

The active principle of the self-amplifying mRNA (saRNA) drug substance is a single-stranded RNA, which self-amplifies upon entering a cell, and the vaccine antigen is translated thereafter. In contrast to uRNA and modRNA that preferably code for a single protein, the

CA 03215771 2023-10-17

coding region of saRNA contains two open reading frames (ORFs). The 5'-ORF encodes the RNA-dependent RNA polymerase such as Venezuelan equine encephalitis virus (VEEV) RNA-dependent RNA polymerase (replicase). The replicase ORF is followed 3' by a subgenomic promoter and a second ORF encoding the antigen. Furthermore, saRNA UTRs contain 5' and 3' conserved sequence elements (CSEs) required for self-amplification. The saRNA contains common structural elements optimized for maximal efficacy of the RNA as the uRNA (5'-cap, 5'-UTR, 3'-UTR, poly(A)-tail). The saRNA preferably contains uridine. The preferred 5' cap structure is beta-S-ARCA(D1) ($m_2^{7,2'-O}$GppSpG).

Cytoplasmic delivery of saRNA initiates an alphavirus-like life cycle. However, the saRNA does not encode for alphaviral structural proteins that are required for genome packaging or cell entry, therefore generation of replication competent viral particles is very unlikely to not possible. Replication does not involve any intermediate steps that generate DNA. The use/uptake of saRNA therefore poses no risk of genomic integration or other permanent genetic modification within the target cell. Furthermore, the saRNA itself prevents its persistent replication by effectively activating innate immune response via recognition of dsRNA intermediates.

Different embodiments of this platform are sequences that are derived from and are modified variants of the following:

RBS004.1 (SEQ ID NO: 24; SEQ ID NO: 7)

Structure beta-S-ARCA(D1)-replicase-S1S2-PP-FI-A30L70

Encoded antigen Viral spike protein (S protein) of the SARS-CoV-2 (S1S2 full-length protein, sequence variant)

RBS004.2 (SEQ ID NO: 25; SEQ ID NO: 7)

Structure beta-S-ARCA(D1)-replicase-S1S2-PP-FI-A30L70

Encoded antigen Viral spike protein (S protein) of the SARS-CoV-2 (S1S2 full-length protein, sequence variant)

CA 03215771 2023-10-17

BNT162c1; RBS004.3 (SEQ ID NO: 26; SEQ ID NO: 5)

Structure beta-S-ARCA(D1)-replicase-RBD-GS-Fibritin-FI-A30L70

Encoded antigen Viral spike protein (S protein) of the SARS-CoV-2 (partial sequence, Receptor Binding Domain (RBD) of S1S2 protein)

RBS004.4 (SEQ ID NO: 27; SEQ ID NO: 28)

Structure beta-S-ARCA(D1)-replicase-RBD-GS-Fibritin-TM-FI-A30L70

Encoded antigen Viral spike protein (S protein) of the SARS-CoV-2 (partial sequence, Receptor Binding Domain (RBD) of S1S2 protein)

Figure 5 schematizes the general structure of the antigen-encoding RNAs.

In some embodiments, vaccine RNA described herein comprises a nucleotide sequence that is a modified variant of a nucleotide sequence selected from the group consisting of SEQ ID NO: 15, 16, 17, 19, 20, 21, 24, 25, 26, 27, 30, and 32. A particularly preferred vaccine RNA described herein comprises a nucleotide sequence that is a modified variant of a nucleotide sequence selected from the group consisting of SEQ ID NO: 15, 17, 19, 21, 25, 26, 30, and 32 such as selected from the group consisting of SEQ ID NO: 17, 19, 21, 26, 30, and 32.

RNA described herein is preferably formulated in lipid nanoparticles (LNP). In one embodiment, the LNP comprise a cationic lipid, a neutral lipid, a steroid, a polymer conjugated lipid; and the RNA. In one embodiment, the cationic lipid is ALC-0315, the neutral lipid is DSPC, the steroid is cholesterol, and the polymer conjugated lipid is ALC-0159. The preferred mode of administration is intramuscular administration, more preferably in aqueous cryoprotectant buffer for intramuscular administration. The drug product is a preferably a preservative-free, sterile dispersion of RNA formulated in lipid nanoparticles (LNP) in aqueous cryoprotectant buffer for intramuscular administration.

CA 03215771 2023-10-17

In different embodiments, the drug product comprises the components shown below, preferably at the proportions or concentrations shown below:

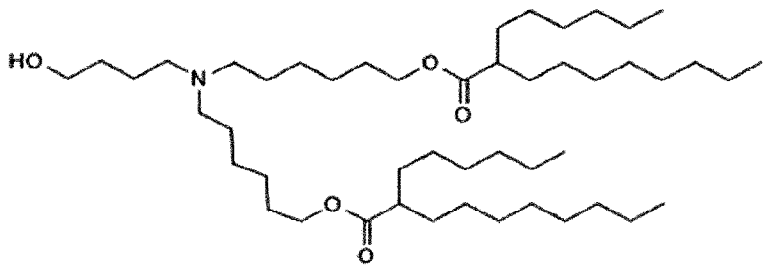
Component	Function	Proportion (mol%)
ALC-0315 [1]	Functional lipid	47.5
ALC-0159 [2]	Functional lipid	1.8
DSPC [3]	Structural lipid	10.0
Cholesterol, synthetic	Structural lipid	40.7

Component	Function	Concentration (mg/mL)
Drug Substance	Active	0.5
ALC-0315 [1]	Functional lipid	7.17
ALC-0159 [2]	Functional lipid	0.89
DSPC [3]	Structural lipid	1.56
Cholesterol, synthetic	Structural lipid	3.1
Sucrose	Cryoprotectant	102.69
NaCl	Buffer	6.0
KCl	Buffer	0.15
Na_2HPO_4	Buffer	1.08
KH_2PO_4	Buffer	0.18
Water for injection	Solvent/Vehicle	q.s.

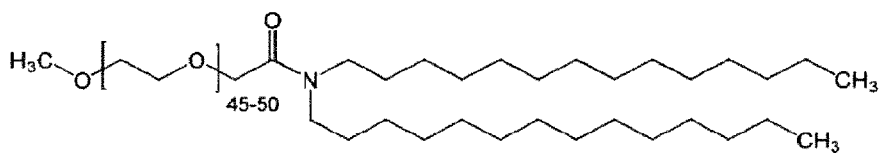
Component	Function	Concentration (mg/mL)
Drug Substance	Active	1.0
ALC-0315 [1]	Functional lipid	13.56
ALC-0159 [2]	Functional lipid	1.77
DSPC [3]	Structural lipid	3.11
Cholesterol, synthetic	Structural lipid	6.20
Sucrose	Cryoprotectant	102.69
NaCl	Buffer	6.0
KCl	Buffer	0.15
Na_2HPO_4	Buffer	1.08
KH_2PO_4	Buffer	0.15
Water for injection	Solvent/Vehicle	q.s.

[1] ALC-0315 = ((4-hydroxybutyl)azanediyl)bis(hexane-6,1-diyl)bis(2-hexyldecanoate) / 6-[N-6-(2-hexyldecanoyloxy)hexyl-N-(4-hydroxybutyl)amino]hexyl 2-hexyldecanoate
[2] ALC-0159 = 2-[(polyethylene glycol)-2000]-*N,N*-ditetradecylacetamide / 2-[2-(ω-methoxy (polyethyleneglycol2000) ethoxy]-N,N-ditetradecylacetamide
[3] DSPC = 1,2-Distearoyl-*sn*-glycero-3-phosphocholine
q.s. = quantum satis (as much as may suffice)

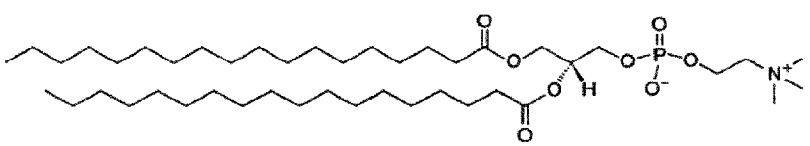
ALC-0315:



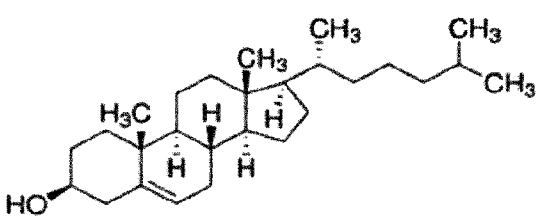
ALC-0159:



DSPC:



Cholesterol:



In one embodiment, the ratio of mRNA to total lipid (N/P) is between 6.0 and 6.5 such as about 6.0 or about 6.3.

Nucleic acid containing particles

Nucleic acids described herein such as RNA encoding a vaccine antigen may be administered formulated as particles.

In the context of the present disclosure, the term "particle" relates to a structured entity formed by molecules or molecule complexes. In one embodiment, the term "particle" relates to a micro- or nano-sized structure, such as a micro- or nano-sized compact structure dispersed in a medium. In one embodiment, a particle is a nucleic acid containing particle such as a particle comprising DNA, RNA or a mixture thereof.

Electrostatic interactions between positively charged molecules such as polymers and lipids and negatively charged nucleic acid are involved in particle formation. This results in complexation and spontaneous formation of nucleic acid particles. In one embodiment, a nucleic acid particle is a nanoparticle.

As used in the present disclosure, "nanoparticle" refers to a particle having an average diameter suitable for parenteral administration.

A "nucleic acid particle" can be used to deliver nucleic acid to a target site of interest (e.g., cell, tissue, organ, and the like). A nucleic acid particle may be formed from at least one cationic or cationically ionizable lipid or lipid-like material, at least one cationic polymer such as protamine, or a mixture thereof and nucleic acid. Nucleic acid particles include lipid nanoparticle (LNP)-based and lipoplex (LPX)-based formulations.

Without intending to be bound by any theory, it is believed that the cationic or cationically ionizable lipid or lipid-like material and/or the cationic polymer combine together with the nucleic acid to form aggregates, and this aggregation results in colloidally stable particles.

In one embodiment, particles described herein further comprise at least one lipid or lipid-like material other than a cationic or cationically ionizable lipid or lipid-like material, at least one polymer other than a cationic polymer, or a mixture thereof

In some embodiments, nucleic acid particles comprise more than one type of nucleic acid molecules, where the molecular parameters of the nucleic acid molecules may be similar or different from each other, like with respect to molar mass or fundamental structural elements such as molecular architecture, capping, coding regions or other features,

Nucleic acid particles described herein may have an average diameter that in one embodiment ranges from about 30 nm to about 1000 nm, from about 50 nm to about 800 nm, from about 70 nm to about 600 nm, from about 90 nm to about 400 nm, or from about 100 nm to about 300 nm.

CA 03215771 2023-10-17

Nucleic acid particles described herein may exhibit a polydispersity index less than about 0.5, less than about 0.4, less than about 0.3, or about 0.2 or less. By way of example, the nucleic acid particles can exhibit a polydispersity index in a range of about 0.1 to about 0.3 or about 0.2 to about 0.3.

With respect to RNA lipid particles, the N/P ratio gives the ratio of the nitrogen groups in the lipid to the number of phosphate groups in the RNA. It is correlated to the charge ratio, as the nitrogen atoms (depending on the pH) are usually positively charged and the phosphate groups are negatively charged. The N/P ratio, where a charge equilibrium exists, depends on the pH. Lipid formulations are frequently formed at N/P ratios larger than four up to twelve, because positively charged nanoparticles are considered favorable for transfection. In that case, RNA is considered to be completely bound to nanoparticles.

Nucleic acid particles described herein can be prepared using a wide range of methods that may involve obtaining a colloid from at least one cationic or cationically ionizable lipid or lipid-like material and/or at least one cationic polymer and mixing the colloid with nucleic acid to obtain nucleic acid particles.

The term "colloid" as used herein relates to a type of homogeneous mixture in which dispersed particles do not settle out. The insoluble particles in the mixture are microscopic, with particle sizes between 1 and 1000 nanometers. The mixture may be termed a colloid or a colloidal suspension. Sometimes the term "colloid" only refers to the particles in the mixture and not the entire suspension.

For the preparation of colloids comprising at least one cationic or cationically ionizable lipid or lipid-like material and/or at least one cationic polymer methods are applicable herein that are conventionally used for preparing liposomal vesicles and are appropriately adapted. The most commonly used methods for preparing liposomal vesicles share the following fundamental stages: (i) lipids dissolution in organic solvents, (ii) drying of the resultant solution, and (iii) hydration of dried lipid (using various aqueous media).

In the film hydration method, lipids are firstly dissolved in a suitable organic solvent, and dried down to yield a thin film at the bottom of the flask. The obtained lipid film is hydrated using an appropriate aqueous medium to produce a liposomal dispersion. Furthermore, an additional downsizing step may be included.

CA 03215771 2023-10-17

Reverse phase evaporation is an alternative method to the film hydration for preparing liposomal vesicles that involves formation of a water-in-oil emulsion between an aqueous phase and an organic phase containing lipids. A brief sonication of this mixture is required for system homogenization. The removal of the organic phase under reduced pressure yields a milky gel that turns subsequently into a liposomal suspension.

The term "ethanol injection technique" refers to a process, in which an ethanol solution comprising lipids is rapidly injected into an aqueous solution through a needle. This action disperses the lipids throughout the solution and promotes lipid structure formation, for example lipid vesicle formation such as liposome formation. Generally, the RNA lipoplex particles described herein are obtainable by adding RNA to a colloidal liposome dispersion. Using the ethanol injection technique, such colloidal liposome dispersion is, in one embodiment, formed as follows: an ethanol solution comprising lipids, such as cationic lipids and additional lipids, is injected into an aqueous solution under stirring. In one embodiment, the RNA lipoplex particles described herein are obtainable without a step of extrusion.

The term "extruding" or "extrusion" refers to the creation of particles having a fixed, cross-sectional profile. In particular, it refers to the downsizing of a particle, whereby the particle is forced through filters with defined pores.

Other methods having organic solvent free characteristics may also be used according to the present disclosure for preparing a colloid.

LNPs typically comprise four components: ionizable cationic lipids, neutral lipids such as phospholipids, a steroid such as cholesterol, and a polymer conjugated lipid such as polyethylene glycol (PEG)-lipids. Each component is responsible for payload protection, and enables effective intracellular delivery. LNPs may be prepared by mixing lipids dissolved in ethanol rapidly with nucleic acid in an aqueous buffer.

The term "average diameter" refers to the mean hydrodynamic diameter of particles as measured by dynamic laser light scattering (DLS) with data analysis using the so-called cumulant algorithm, which provides as results the so-called $Z_{average}$ with the dimension of a length, and the polydispersity index (PI), which is dimensionless (Koppel, D., J. Chem. Phys. 57, 1972, pp 4814-4820, ISO 13321). Here "average diameter", "diameter" or "size" for particles is used synonymously with this value of the $Z_{average}$.

CA 03215771 2023-10-17

The "polydispersity index" is preferably calculated based on dynamic light scattering measurements by the so-called cumulant analysis as mentioned in the definition of the "average diameter". Under certain prerequisites, it can be taken as a measure of the size distribution of an ensemble of nanoparticles.

Different types of nucleic acid containing particles have been described previously to be suitable for delivery of nucleic acid in particulate form (e.g. Kaczmarek, J. C. et al., 2017, Genome Medicine 9, 60). For non-viral nucleic acid delivery vehicles, nanoparticle encapsulation of nucleic acid physically protects nucleic acid from degradation and, depending on the specific chemistry, can aid in cellular uptake and endosomal escape.

The present disclosure describes particles comprising nucleic acid, at least one cationic or cationically ionizable lipid or lipid-like material, and/or at least one cationic polymer which associate with nucleic acid to form nucleic acid particles and compositions comprising such particles. The nucleic acid particles may comprise nucleic acid which is complexed in different forms by non-covalent interactions to the particle. The particles described herein are not viral particles, in particular infectious viral particles, i.e., they are not able to virally infect cells.

Suitable cationic or cationically ionizable lipids or lipid-like materials and cationic polymers are those that form nucleic acid particles and are included by the term "particle forming components" or "particle forming agents". The term "particle forming components" or "particle forming agents" relates to any components which associate with nucleic acid to form nucleic acid particles. Such components include any component which can be part of nucleic acid particles.

Some embodiments described herein relate to compositions, methods and uses involving more than one, e.g., 2, 3, 4, 5, 6 or even more nucleic acid species such as RNA species, e.g., a) a nucleic acid comprising a first nucleotide sequence encoding an amino acid sequence comprising at least a fragment of a parental virus protein, wherein amino acid positions in the at least a fragment of a parental virus protein are modified to comprise amino acids found in the corresponding amino acid positions of one or more virus protein variants; and b) a nucleic acid comprising a second nucleotide sequence encoding an amino acid sequence comprising at least a fragment of a parental virus protein, wherein amino acid positions in the at least a

CA 03215771 2023-10-17

fragment of a parental virus protein are modified to comprise amino acids found in the corresponding amino acid positions of one or more virus protein variants.

In some embodiments, a set of nucleic acids is provided, each of which includes a nucleotide sequence encoding an amino acid sequence comprising at least a fragment of a parental virus protein, wherein each encoded fragment of the set is substantially identical to other encoded fragments of the set, and to the parental encoded fragment, except that modification(s) relative to the parental encoded fragment that have arisen in circulating viral variants are present in encoded fragments of the set.

In a particulate formulation, it is possible that each nucleic acid species is separately formulated as an individual particulate formulation. In that case, each individual particulate formulation will comprise one nucleic acid species. The individual particulate formulations may be present as separate entities, e.g. in separate containers. Such formulations are obtainable by providing each nucleic acid species separately (typically each in the form of a nucleic acid-containing solution) together with a particle-forming agent, thereby allowing the formation of particles. Respective particles will contain exclusively the specific nucleic acid species that is being provided when the particles are formed (individual particulate formulations).

In one embodiment, a composition such as a pharmaceutical composition comprises more than one individual particle formulation. Respective pharmaceutical compositions are referred to as mixed particulate formulations. Mixed particulate formulations according to the invention are obtainable by forming, separately, individual particulate formulations, as described above, followed by a step of mixing of the individual particulate formulations. By the step of mixing, a formulation comprising a mixed population of nucleic acid-containing particles is obtainable. Individual particulate populations may be together in one container, comprising a mixed population of individual particulate formulations.

Alternatively, it is possible that different nucleic acid species are formulated together as a combined particulate formulation. Such formulations are obtainable by providing a combined formulation (typically combined solution) of different RNA species together with a particle-forming agent, thereby allowing the formation of particles. As opposed to a mixed particulate formulation, a combined particulate formulation will typically comprise particles which

CA 03215771 2023-10-17

comprise more than one RNA species. In a combined particulate composition different RNA species are typically present together in a single particle.

Cationic polymer

Given their high degree of chemical flexibility, polymers are commonly used materials for nanoparticle-based delivery. Typically, cationic polymers are used to electrostatically condense the negatively charged nucleic acid into nanoparticles. These positively charged groups often consist of amines that change their state of protonation in the pH range between 5.5 and 7.5, thought to lead to an ion imbalance that results in endosomal rupture. Polymers such as poly-L-lysine, polyamidoamine, protamine and polyethyleneimine, as well as naturally occurring polymers such as chitosan have all been applied to nucleic acid delivery and are suitable as cationic polymers herein. In addition, some investigators have synthesized polymers specifically for nucleic acid delivery. Poly(β-amino esters), in particular, have gained widespread use in nucleic acid delivery owing to their ease of synthesis and biodegradability. Such synthetic polymers are also suitable as cationic polymers herein.

A "polymer," as used herein, is given its ordinary meaning, i.e., a molecular structure comprising one or more repeat units (monomers), connected by covalent bonds. The repeat units can all be identical, or in some cases, there can be more than one type of repeat unit present within the polymer. In some cases, the polymer is biologically derived, i.e., a biopolymer such as a protein. In some cases, additional moieties can also be present in the polymer, for example targeting moieties such as those described herein.

If more than one type of repeat unit is present within the polymer, then the polymer is said to be a "copolymer." It is to be understood that the polymer being employed herein can be a copolymer. The repeat units forming the copolymer can be arranged in any fashion. For example, the repeat units can be arranged in a random order, in an alternating order, or as a "block" copolymer, i.e., comprising one or more regions each comprising a first repeat unit (e.g., a first block), and one or more regions each comprising a second repeat unit (e.g., a second block), etc. Block copolymers can have two (a diblock copolymer), three (a triblock copolymer), or more numbers of distinct blocks.

CA 03215771 2023-10-17

In certain embodiments, the polymer is biocompatible. Biocompatible polymers are polymers that typically do not result in significant cell death at moderate concentrations. In certain embodiments, the biocompatible polymer is biodegradable, i.e., the polymer is able to degrade, chemically and/or biologically, within a physiological environment, such as within the body.

In certain embodiments, polymer may be protamine or polyalkyleneimine, in particular protamine.

The term "protamine" refers to any of various strongly basic proteins of relatively low molecular weight that are rich in arginine and are found associated especially with DNA in place of somatic histones in the sperm cells of various animals (as fish). In particular, the term "protamine" refers to proteins found in fish sperm that are strongly basic, are soluble in water, are not coagulated by heat, and yield chiefly arginine upon hydrolysis. In purified form, they are used in a long-acting formulation of insulin and to neutralize the anticoagulant effects of heparin.

According to the disclosure, the term "protamine" as used herein is meant to comprise any protamine amino acid sequence obtained or derived from natural or biological sources including fragments thereof and multimeric forms of said amino acid sequence or fragment thereof as well as (synthesized) polypeptides which are artificial and specifically designed for specific purposes and cannot be isolated from native or biological sources.

In one embodiment, the polyalkyleneimine comprises polyethylenimine and/or polypropylenimine, preferably polyethyleneimine. A preferred polyalkyleneimine is polyethyleneimine (PEI). The average molecular weight of PEI is preferably $0.75 \cdot 10^2$ to 10^7 Da, preferably 1000 to 10^5 Da, more preferably 10000 to 40000 Da, more preferably 15000 to 30000 Da, even more preferably 20000 to 25000 Da.

Preferred according to the disclosure is linear polyalkyleneimine such as linear polyethyleneimine (PEI).

Cationic polymers (including polycationic polymers) contemplated for use herein include any cationic polymers which are able to electrostatically bind nucleic acid. In one embodiment, cationic polymers contemplated for use herein include any cationic polymers with which

CA 03215771 2023-10-17

nucleic acid can be associated, e.g. by forming complexes with the nucleic acid or forming vesicles in which the nucleic acid is enclosed or encapsulated.

Particles described herein may also comprise polymers other than cationic polymers, i.e., non-cationic polymers and/or anionic polymers. Collectively, anionic and neutral polymers are referred to herein as non-cationic polymers.

Lipid and lipid-like material

The terms "lipid" and "lipid-like material" are broadly defined herein as molecules which comprise one or more hydrophobic moieties or groups and optionally also one or more hydrophilic moieties or groups. Molecules comprising hydrophobic moieties and hydrophilic moieties are also frequently denoted as amphiphiles. Lipids are usually poorly soluble in water. In an aqueous environment, the amphiphilic nature allows the molecules to self-assemble into organized structures and different phases. One of those phases consists of lipid bilayers, as they are present in vesicles, multilamellar/unilamellar liposomes, or membranes in an aqueous environment. Hydrophobicity can be conferred by the inclusion of apolar groups that include, but are not limited to, long-chain saturated and unsaturated aliphatic hydrocarbon groups and such groups substituted by one or more aromatic, cycloaliphatic, or heterocyclic group(s). The hydrophilic groups may comprise polar and/or charged groups and include carbohydrates, phosphate, carboxylic, sulfate, amino, sulfhydryl, nitro, hydroxyl, and other like groups.

As used herein, the term "amphiphilic" refers to a molecule having both a polar portion and a non-polar portion. Often, an amphiphilic compound has a polar head attached to a long hydrophobic tail. In some embodiments, the polar portion is soluble in water, while the non-polar portion is insoluble in water. In addition, the polar portion may have either a formal positive charge, or a formal negative charge. Alternatively, the polar portion may have both a formal positive and a negative charge, and be a zwitterion or inner salt. For purposes of the disclosure, the amphiphilic compound can be, but is not limited to, one or a plurality of natural or non-natural lipids and lipid-like compounds.

The term "lipid-like material", "lipid-like compound" or "lipid-like molecule" relates to substances that structurally and/or functionally relate to lipids but may not be considered as

CA 03215771 2023-10-17

lipids in a strict sense. For example, the term includes compounds that are able to form amphiphilic layers as they are present in vesicles, multilamellar/unilamellar liposomes, or membranes in an aqueous environment and includes surfactants, or synthesized compounds with both hydrophilic and hydrophobic moieties. Generally speaking, the term refers to molecules, which comprise hydrophilic and hydrophobic moieties with different structural organization, which may or may not be similar to that of lipids. As used herein, the term "lipid" is to be construed to cover both lipids and lipid-like materials unless otherwise indicated herein or clearly contradicted by context.

Specific examples of amphiphilic compounds that may be included in an amphiphilic layer include, but are not limited to, phospholipids, aminolipids and sphingolipids.

In certain embodiments, the amphiphilic compound is a lipid. The term "lipid" refers to a group of organic compounds that are characterized by being insoluble in water, but soluble in many organic solvents. Generally, lipids may be divided into eight categories: fatty acids, glycerolipids, glycerophospholipids, sphingolipids, saccharolipids, polyketides (derived from condensation of ketoacyl subunits), sterol lipids and prenol lipids (derived from condensation of isoprene subunits). Although the term "lipid" is sometimes used as a synonym for fats, fats are a subgroup of lipids called triglycerides. Lipids also encompass molecules such as fatty acids and their derivatives (including tri-, di-, monoglycerides, and phospholipids), as well as sterol-containing metabolites such as cholesterol.

Fatty acids, or fatty acid residues are a diverse group of molecules made of a hydrocarbon chain that terminates with a carboxylic acid group; this arrangement confers the molecule with a polar, hydrophilic end, and a nonpolar, hydrophobic end that is insoluble in water. The carbon chain, typically between four and 24 carbons long, may be saturated or unsaturated, and may be attached to functional groups containing oxygen, halogens, nitrogen, and sulfur. If a fatty acid contains a double bond, there is the possibility of either a cis or trans geometric isomerism, which significantly affects the molecule's configuration. Cis-double bonds cause the fatty acid chain to bend, an effect that is compounded with more double bonds in the chain. Other major lipid classes in the fatty acid category are the fatty esters and fatty amides. Glycerolipids are composed of mono-, di-, and tri-substituted glycerols, the best-known being the fatty acid triesters of glycerol, called triglycerides. The word "triacylglycerol" is sometimes

CA 03215771 2023-10-17

used synonymously with "triglyceride". In these compounds, the three hydroxyl groups of glycerol are each esterified, typically by different fatty acids. Additional subclasses of glycerolipids are represented by glycosylglycerols, which are characterized by the presence of one or more sugar residues attached to glycerol via a glycosidic linkage.

The glycerophospholipids are amphipathic molecules (containing both hydrophobic and hydrophilic regions) that contain a glycerol core linked to two fatty acid-derived "tails" by ester linkages and to one "head" group by a phosphate ester linkage. Examples of glycerophospholipids, usually referred to as phospholipids (though sphingomyelins are also classified as phospholipids) are phosphatidylcholine (also known as PC, GPCho or lecithin), phosphatidylethanolamine (PE or GPEtn) and phosphatidylserine (PS or GPSer).

Sphingolipids are a complex family of compounds that share a common structural feature, a sphingoid base backbone. The major sphingoid base in mammals is commonly referred to as sphingosine. Ceramides (N-acyl-sphingoid bases) are a major subclass of sphingoid base derivatives with an amide-linked fatty acid. The fatty acids are typically saturated or mono-unsaturated with chain lengths from 16 to 26 carbon atoms. The major phosphosphingolipids of mammals are sphingomyelins (ceramide phosphocholines), whereas insects contain mainly ceramide phosphoethanolamines and fungi have phytoceramide phosphoinositols and mannose-containing headgroups. The glycosphingolipids are a diverse family of molecules composed of one or more sugar residues linked via a glycosidic bond to the sphingoid base. Examples of these are the simple and complex glycosphingolipids such as cerebrosides and gangliosides.

Sterol lipids, such as cholesterol and its derivatives, or tocopherol and its derivatives, are an important component of membrane lipids, along with the glycerophospholipids and sphingomyelins.

Saccharolipids describe compounds in which fatty acids are linked directly to a sugar backbone, forming structures that are compatible with membrane bilayers. In the saccharolipids, a monosaccharide substitutes for the glycerol backbone present in glycerolipids and glycerophospholipids. The most familiar saccharolipids are the acylated glucosamine precursors of the Lipid A component of the lipopolysaccharides in Gram-negative bacteria. Typical lipid A molecules are disaccharides of glucosamine, which are derivatized

CA 03215771 2023-10-17

with as many as seven fatty-acyl chains. The minimal lipopolysaccharide required for growth in E. coli is Kdo2-Lipid A, a hexa-acylated disaccharide of glucosamine that is glycosylated with two 3-deoxy-D-manno-octulosonic acid (Kdo) residues.

Polyketides are synthesized by polymerization of acetyl and propionyl subunits by classic enzymes as well as iterative and multimodular enzymes that share mechanistic features with the fatty acid synthases. They comprise a large number of secondary metabolites and natural products from animal, plant, bacterial, fungal and marine sources, and have great structural diversity. Many polyketides are cyclic molecules whose backbones are often further modified by glycosylation, methylation, hydroxylation, oxidation, or other processes.

According to the disclosure, lipids and lipid-like materials may be cationic, anionic or neutral. Neutral lipids or lipid-like materials exist in an uncharged or neutral zwitterionic form at a selected pH.

Cationic or cationically ionizable lipids or lipid-like materials

The nucleic acid particles described herein may comprise at least one cationic or cationically ionizable lipid or lipid-like material as particle forming agent. Cationic or cationically ionizable lipids or lipid-like materials contemplated for use herein include any cationic or cationically ionizable lipids or lipid-like materials which are able to electrostatically bind nucleic acid. In one embodiment, cationic or cationically ionizable lipids or lipid-like materials contemplated for use herein can be associated with nucleic acid, e.g. by forming complexes with the nucleic acid or forming vesicles in which the nucleic acid is enclosed or encapsulated.

As used herein, a "cationic lipid" or "cationic lipid-like material" refers to a lipid or lipid-like material having a net positive charge. Cationic lipids or lipid-like materials bind negatively charged nucleic acid by electrostatic interaction. Generally, cationic lipids possess a lipophilic moiety, such as a sterol, an acyl chain, a diacyl or more acyl chains, and the head group of the lipid typically carries the positive charge.

In certain embodiments, a cationic lipid or lipid-like material has a net positive charge only at certain pH, in particular acidic pH, while it has preferably no net positive charge, preferably has no charge, i.e., it is neutral, at a different, preferably higher pH such as physiological pH.

This ionizable behavior is thought to enhance efficacy through helping with endosomal escape and reducing toxicity as compared with particles that remain cationic at physiological pH.

For purposes of the present disclosure, such "cationically ionizable" lipids or lipid-like materials are comprised by the term "cationic lipid or lipid-like material" unless contradicted by the circumstances.

In one embodiment, the cationic or cationically ionizable lipid or lipid-like material comprises a head group which includes at least one nitrogen atom (N) which is positive charged or capable of being protonated.

Examples of cationic lipids include, but are not limited to 1,2-dioleoyl-3-trimethylammonium propane (DOTAP); N,N-dimethyl-2,3-dioleyloxypropylamine (DODMA), 1,2-di-O-octadecenyl-3-trimethylammonium propane (DOTMA), 3-(N—(N',N'-dimethylaminoethane)-carbamoyl)cholesterol (DC-Chol), dimethyldioctadecylammonium (DDAB); 1,2-dioleoyl-3-dimethylammonium-propane (DODAP); 1,2-diacyloxy-3-dimethylammonium propanes; 1,2-dialkyloxy-3-dimethylammonium propanes; dioctadecyldimethyl ammonium chloride (DODAC), 1,2-distearyloxy-N,N-dimethyl-3-aminopropane (DSDMA), 2,3-di(tetradecoxy)propyl-(2-hydroxyethyl)-dimethylazanium (DMRIE), 1,2-dimyristoyl-sn-glycero-3-ethylphosphocholine (DMEPC), l,2-dimyristoyl-3-trimethylammonium propane (DMTAP), 1,2-dioleyloxypropyl-3-dimethyl-hydroxyethyl ammonium bromide (DORIE), and 2,3-dioleoyloxy- N-[2(spermine carboxamide)ethyl]-N,N-dimethyl-l-propanamium trifluoroacetate (DOSPA), 1,2-dilinoleyloxy-N,N-dimethylaminopropane (DLinDMA), 1,2-dilinolenyloxy-N,N-dimethylaminopropane (DLenDMA), dioctadecylamidoglycyl spermine (DOGS), 3-dimethylamino-2-(cholest-5-en-3-beta-oxybutan-4-oxy)-1-(cis,cis-9,12-octadecadienoxy)propane (CLinDMA), 2-[5'-(cholest-5-en-3-beta-oxy)-3'-oxapentoxy)-3-dimethyl-1-(cis,cis-9',12'-octadecadienoxy)propane (CpLinDMA), N,N-dimethyl-3,4-dioleyloxybenzylamine (DMOBA), 1,2-N,N'-dioleylcarbamyl-3-dimethylaminopropane (DOcarbDAP), 2,3-Dilinoleoyloxy-N,N-dimethylpropylamine (DLinDAP), 1,2-N,N'-Dilinoleylcarbamyl-3-dimethylaminopropane (DLincarbDAP), 1,2-Dilinoleoylcarbamyl-3-dimethylaminopropane (DLinCDAP), 2,2-dilinoleyl-4-dimethylaminomethyl-[1,3]-dioxolane (DLin-K-DMA), 2,2-dilinoleyl-4-dimethylaminoethyl-[1,3]-dioxolane (DLin-K-XTC2-DMA), 2,2-dilinoleyl-4-(2-dimethylaminoethyl)-[1,3]-dioxolane (DLin-KC2-DMA), heptatriaconta-

CA 03215771 2023-10-17

6,9,28,31-tetraen-19-yl-4-(dimethylamino)butanoate (DLin-MC3-DMA), N-(2-Hydroxyethyl)-N,N-dimethyl-2,3-bis(tetradecyloxy)-1-propanaminium bromide (DMRIE), (±)-N-(3-aminopropyl)-N,N-dimethyl-2,3-bis(cis-9-tetradecenyloxy)-1-propanaminium bromide (GAP-DMORIE), (±)-N-(3-aminopropyl)-N,N-dimethyl-2,3-bis(dodecyloxy)-1-propanaminium bromide (GAP-DLRIE), (±)-N-(3-aminopropyl)-N,N-dimethyl-2,3-bis(tetradecyloxy)-1-propanaminium bromide (GAP-DMRIE), N-(2-Aminoethyl)-N,N-dimethyl-2,3-bis(tetradecyloxy)-1-propanaminium bromide (βAE-DMRIE), N-(4-carboxybenzyl)-N,N-dimethyl-2,3-bis(oleoyloxy)propan-1-aminium (DOBAQ), 2-({8-[(3β)-cholest-5-en-3-yloxy]octyl}oxy)-N,N-dimethyl-3-[(9Z,12Z)-octadeca-9,12-dien-1-yloxy]propan-1-amine (Octyl-CLinDMA), 1,2-dimyristoyl-3-dimethylammonium-propane (DMDAP), 1,2-dipalmitoyl-3-dimethylammonium-propane (DPDAP), N1-[2-((1S)-1-[(3-aminopropyl)amino]-4-[di(3-amino-propyl)amino]butylcarboxamido)ethyl]-3,4-di[oleyloxy]-benzamide (MVL5), 1,2-dioleoyl-sn-glycero-3-ethylphosphocholine (DOEPC), 2,3-bis(dodecyloxy)-N-(2-hydroxyethyl)-N,N-dimethylpropan-1-amonium bromide (DLRIE), N-(2-aminoethyl)-N,N-dimethyl-2,3-bis(tetradecyloxy)propan-1-aminium bromide (DMORIE), di((Z)-non-2-en-1-yl) 8,8'-((((2(dimethylamino)ethyl)thio)carbonyl)azanediyl)dioctanoate (ATX), N,N-dimethyl-2,3-bis(dodecyloxy)propan-1-amine (DLDMA), N,N-dimethyl-2,3-bis(tetradecyloxy)propan-1-amine (DMDMA), Di((Z)-non-2-en-1-yl)-9-((4-(dimethylaminobutanoyl)oxy)heptadecanedioate (L319), N-Dodecyl-3-((2-dodecylcarbamoyl-ethyl)-{2-[(2-dodecylcarbamoyl-ethyl)-2-{(2-dodecylcarbamoyl-ethyl)-[2-(2-dodecylcarbamoyl-ethylamino)-ethyl]-amino}-ethylamino)propionamide (lipidoid 98N_{12}-5), 1-[2-[bis(2-hydroxydodecyl)amino]ethyl-[2-[4-[2-[bis(2 hydroxydodecyl)amino]ethyl]piperazin-1-yl]ethyl]amino]dodecan-2-ol (lipidoid C12-200).

In some embodiments, the cationic lipid may comprise from about 10 mol % to about 100 mol %, about 20 mol % to about 100 mol %, about 30 mol % to about 100 mol %, about 40 mol % to about 100 mol %, or about 50 mol % to about 100 mol % of the total lipid present in the particle.

CA 03215771 2023-10-17

Additional lipids or lipid-like materials

Particles described herein may also comprise lipids or lipid-like materials other than cationic or cationically ionizable lipids or lipid-like materials, i.e., non-cationic lipids or lipid-like materials (including non-cationically ionizable lipids or lipid-like materials). Collectively, anionic and neutral lipids or lipid-like materials are referred to herein as non-cationic lipids or lipid-like materials. Optimizing the formulation of nucleic acid particles by addition of other hydrophobic moieties, such as cholesterol and lipids, in addition to an ionizable/cationic lipid or lipid-like material may enhance particle stability and efficacy of nucleic acid delivery.

An additional lipid or lipid-like material may be incorporated which may or may not affect the overall charge of the nucleic acid particles. In certain embodiments, the additional lipid or lipid-like material is a non-cationic lipid or lipid-like material. The non-cationic lipid may comprise, e.g., one or more anionic lipids and/or neutral lipids. As used herein, an "anionic lipid" refers to any lipid that is negatively charged at a selected pH. As used herein, a "neutral lipid" refers to any of a number of lipid species that exist either in an uncharged or neutral zwitterionic form at a selected pH. In preferred embodiments, the additional lipid comprises one of the following neutral lipid components: (1) a phospholipid, (2) cholesterol or a derivative thereof; or (3) a mixture of a phospholipid and cholesterol or a derivative thereof. Examples of cholesterol derivatives include, but are not limited to, cholestanol, cholestanone, cholestenone, coprostanol, cholesteryl-2'-hydroxyethyl ether, cholesteryl-4'- hydroxybutyl ether, tocopherol and derivatives thereof, and mixtures thereof.

Specific phospholipids that can be used include, but are not limited to, phosphatidylcholines, phosphatidylethanolamines, phosphatidylglycerols, phosphatidic acids, phosphatidylserines or sphingomyelin. Such phospholipids include in particular diacylphosphatidylcholines, such as distearoylphosphatidylcholine (DSPC), dioleoylphosphatidylcholine (DOPC), dimyristoylphosphatidylcholine (DMPC), dipentadecanoylphosphatidylcholine, dilauroylphosphatidylcholine, dipalmitoylphosphatidylcholine (DPPC), diarachidoylphosphatidylcholine (DAPC), dibehenoylphosphatidylcholine (DBPC), ditricosanoylphosphatidylcholine (DTPC), dilignoceroylphatidylcholine (DLPC), palmitoyloleoyl-phosphatidylcholine (POPC), 1,2-di-O-octadecenyl-sn-glycero-3-phosphocholine (18:0 Diether PC), 1-oleoyl-2-cholesterylhemisuccinoyl-sn-glycero-3-

CA 03215771 2023-10-17

phosphocholine (OChemsPC), 1-hexadecyl-sn-glycero-3-phosphocholine (C16 Lyso PC) and phosphatidylethanolamines, in particular diacylphosphatidylethanolamines, such as dioleoylphosphatidylethanolamine (DOPE), distearoyl-phosphatidylethanolamine (DSPE), dipalmitoyl-phosphatidylethanolamine (DPPE), dimyristoyl-phosphatidylethanolamine (DMPE), dilauroyl-phosphatidylethanolamine (DLPE), diphytanoyl-phosphatidylethanolamine (DPyPE), and further phosphatidylethanolamine lipids with different hydrophobic chains.

In certain preferred embodiments, the additional lipid is DSPC or DSPC and cholesterol.

In certain embodiments, the nucleic acid particles include both a cationic lipid and an additional lipid.

In one embodiment, particles described herein include a polymer conjugated lipid such as a pegylated lipid. The term "pegylated lipid" refers to a molecule comprising both a lipid portion and a polyethylene glycol portion. Pegylated lipids are known in the art.

Without wishing to be bound by theory, the amount of the at least one cationic lipid compared to the amount of the at least one additional lipid may affect important nucleic acid particle characteristics, such as charge, particle size, stability, tissue selectivity, and bioactivity of the nucleic acid. Accordingly, in some embodiments, the molar ratio of the at least one cationic lipid to the at least one additional lipid is from about 10:0 to about 1:9, about 4:1 to about 1:2, or about 3:1 to about 1:1.

In some embodiments, the non-cationic lipid, in particular neutral lipid, (e.g., one or more phospholipids and/or cholesterol) may comprise from about 0 mol % to about 90 mol %, from about 0 mol % to about 80 mol %, from about 0 mol % to about 70 mol %, from about 0 mol % to about 60 mol %, or from about 0 mol % to about 50 mol %, of the total lipid present in the particle.

Lipoplex Particles

In certain embodiments of the present disclosure, the RNA described herein may be present in RNA lipoplex particles.

In the context of the present disclosure, the term "RNA lipoplex particle" relates to a particle that contains lipid, in particular cationic lipid, and RNA. Electrostatic interactions between positively charged liposomes and negatively charged RNA results in complexation and

spontaneous formation of RNA lipoplex particles. Positively charged liposomes may be generally synthesized using a cationic lipid, such as DOTMA, and additional lipids, such as DOPE. In one embodiment, a RNA lipoplex particle is a nanoparticle.

In certain embodiments, the RNA lipoplex particles include both a cationic lipid and an additional lipid. In an exemplary embodiment, the cationic lipid is DOTMA and the additional lipid is DOPE.

In some embodiments, the molar ratio of the at least one cationic lipid to the at least one additional lipid is from about 10:0 to about 1:9, about 4:1 to about 1:2, or about 3:1 to about 1:1. In specific embodiments, the molar ratio may be about 3:1, about 2.75:1, about 2.5:1, about 2.25:1, about 2:1, about 1.75:1, about 1.5:1, about 1.25:1, or about 1:1. In an exemplary embodiment, the molar ratio of the at least one cationic lipid to the at least one additional lipid is about 2:1.

RNA lipoplex particles described herein have an average diameter that in one embodiment ranges from about 200 nm to about 1000 nm, from about 200 nm to about 800 nm, from about 250 to about 700 nm, from about 400 to about 600 nm, from about 300 nm to about 500 nm, or from about 350 nm to about 400 nm. In specific embodiments, the RNA lipoplex particles have an average diameter of about 200 nm, about 225 nm, about 250 nm, about 275 nm, about 300 nm, about 325 nm, about 350 nm, about 375 nm, about 400 nm, about 425 nm, about 450 nm, about 475 nm, about 500 nm, about 525 nm, about 550 nm, about 575 nm, about 600 nm, about 625 nm, about 650 nm, about 700 nm, about 725 nm, about 750 nm, about 775 nm, about 800 nm, about 825 nm, about 850 nm, about 875 nm, about 900 nm, about 925 nm, about 950 nm, about 975 nm, or about 1000 nm. In an embodiment, the RNA lipoplex particles have an average diameter that ranges from about 250 nm to about 700 nm. In another embodiment, the RNA lipoplex particles have an average diameter that ranges from about 300 nm to about 500 nm. In an exemplary embodiment, the RNA lipoplex particles have an average diameter of about 400 nm.

The RNA lipoplex particles and compositions comprising RNA lipoplex particles described herein are useful for delivery of RNA to a target tissue after parenteral administration, in particular after intravenous administration. The RNA lipoplex particles may be prepared using liposomes that may be obtained by injecting a solution of the lipids in ethanol into water or a

CA 03215771 2023-10-17

suitable aqueous phase. In one embodiment, the aqueous phase has an acidic pH. In one embodiment, the aqueous phase comprises acetic acid, e.g., in an amount of about 5 mM. Liposomes may be used for preparing RNA lipoplex particles by mixing the liposomes with RNA. In one embodiment, the liposomes and RNA lipoplex particles comprise at least one cationic lipid and at least one additional lipid. In one embodiment, the at least one cationic lipid comprises 1,2-di-O-octadecenyl-3-trimethylammonium propane (DOTMA) and/or 1,2-dioleoyl-3-trimethylammonium-propane (DOTAP). In one embodiment, the at least one additional lipid comprises 1,2-di-(9Z-octadecenoyl)-sn-glycero-3-phosphoethanolamine (DOPE), cholesterol (Chol) and/or 1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPC). In one embodiment, the at least one cationic lipid comprises 1,2-di-O-octadecenyl-3-trimethylammonium propane (DOTMA) and the at least one additional lipid comprises 1,2-di-(9Z-octadecenoyl)-sn-glycero-3-phosphoethanolamine (DOPE). In one embodiment, the liposomes and RNA lipoplex particles comprise 1,2-di-O-octadecenyl-3-trimethylammonium propane (DOTMA) and 1,2-di-(9Z-octadecenoyl)-sn-glycero-3-phosphoethanolamine (DOPE). Spleen targeting RNA lipoplex particles are described in WO 2013/143683, herein incorporated by reference. It has been found that RNA lipoplex particles having a net negative charge may be used to preferentially target spleen tissue or spleen cells such as antigen-presenting cells, in particular dendritic cells. Accordingly, following administration of the RNA lipoplex particles, RNA accumulation and/or RNA expression in the spleen occurs. Thus, RNA lipoplex particles of the disclosure may be used for expressing RNA in the spleen. In an embodiment, after administration of the RNA lipoplex particles, no or essentially no RNA accumulation and/or RNA expression in the lung and/or liver occurs. In one embodiment, after administration of the RNA lipoplex particles, RNA accumulation and/or RNA expression in antigen presenting cells, such as professional antigen presenting cells in the spleen occurs. Thus, RNA lipoplex particles of the disclosure may be used for expressing RNA in such antigen presenting cells. In one embodiment, the antigen presenting cells are dendritic cells and/or macrophages.

Lipid nanoparticles (LNPs)

In one embodiment, nucleic acid such as RNA described herein is administered in the form of lipid nanoparticles (LNPs). The LNP may comprise any lipid capable of forming a particle to which the one or more nucleic acid molecules are attached, or in which the one or more nucleic acid molecules are encapsulated.

In one embodiment, the LNP comprises one or more cationic lipids, and one or more stabilizing lipids. Stabilizing lipids include neutral lipids and pegylated lipids.

In one embodiment, the LNP comprises a cationic lipid, a neutral lipid, a steroid, a polymer conjugated lipid; and the RNA, encapsulated within or associated with the lipid nanoparticle.

In one embodiment, the LNP comprises from 40 to 55 mol percent, from 40 to 50 mol percent, from 41 to 49 mol percent, from 41 to 48 mol percent, from 42 to 48 mol percent, from 43 to 48 mol percent, from 44 to 48 mol percent, from 45 to 48 mol percent, from 46 to 48 mol percent, from 47 to 48 mol percent, or from 47.2 to 47.8 mol percent of the cationic lipid. In one embodiment, the LNP comprises about 47.0, 47.1, 47.2, 47.3, 47.4, 47.5, 47.6, 47.7, 47.8, 47.9 or 48.0 mol percent of the cationic lipid.

In one embodiment, the neutral lipid is present in a concentration ranging from 5 to 15 mol percent, from 7 to 13 mol percent, or from 9 to 11 mol percent. In one embodiment, the neutral lipid is present in a concentration of about 9.5, 10 or 10.5 mol percent.

In one embodiment, the steroid is present in a concentration ranging from 30 to 50 mol percent, from 35 to 45 mol percent or from 38 to 43 mol percent. In one embodiment, the steroid is present in a concentration of about 40, 41, 42, 43, 44, 45 or 46 mol percent.

In one embodiment, the LNP comprises from 1 to 10 mol percent, from 1 to 5 mol percent, or from 1 to 2.5 mol percent of the polymer conjugated lipid.

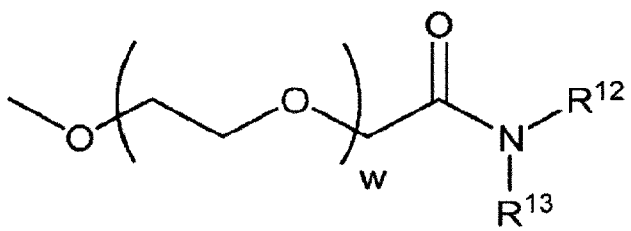
In one embodiment, the LNP comprises from 40 to 50 mol percent a cationic lipid; from 5 to 15 mol percent of a neutral lipid; from 35 to 45 mol percent of a steroid; from 1 to 10 mol percent of a polymer conjugated lipid; and the RNA, encapsulated within or associated with the lipid nanoparticle.

In one embodiment, the mol percent is determined based on total mol of lipid present in the lipid nanoparticle.

In one embodiment, the neutral lipid is selected from the group consisting of DSPC, DPPC, DMPC, DOPC, POPC, DOPE, DOPG, DPPG, POPE, DPPE, DMPE, DSPE, and SM. In one embodiment, the neutral lipid is selected from the group consisting of DSPC, DPPC, DMPC, DOPC, POPC, DOPE and SM. In one embodiment, the neutral lipid is DSPC.

In one embodiment, the steroid is cholesterol.

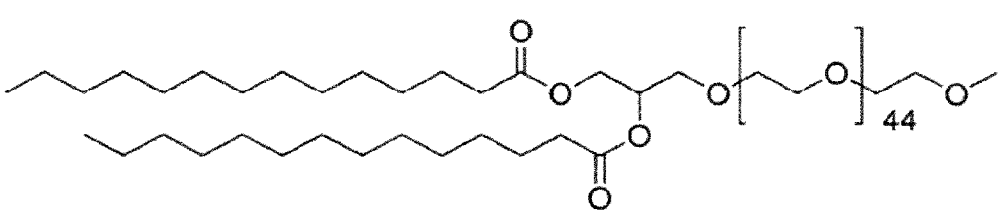
In one embodiment, the polymer conjugated lipid is a pegylated lipid. In one embodiment, the pegylated lipid has the following structure:



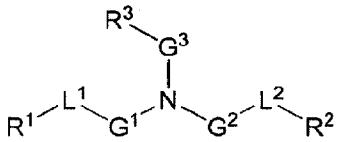
or a pharmaceutically acceptable salt, tautomer or stereoisomer thereof, wherein:

R^{12} and R^{13} are each independently a straight or branched, saturated or unsaturated alkyl chain containing from 10 to 30 carbon atoms, wherein the alkyl chain is optionally interrupted by one or more ester bonds; and w has a mean value ranging from 30 to 60. In one embodiment, R^{12} and R^{13} are each independently straight, saturated alkyl chains containing from 12 to 16 carbon atoms. In one embodiment, w has a mean value ranging from 40 to 55. In one embodiment, the average w is about 45. In one embodiment, R^{12} and R^{13} are each independently a straight, saturated alkyl chain containing about 14 carbon atoms, and w has a mean value of about 45.

In one embodiment, the pegylated lipid is DMG-PEG 2000, e.g., having the following structure:



In some embodiments, the cationic lipid component of the LNPs has the structure of Formula (III):



(III)

or a pharmaceutically acceptable salt, tautomer, prodrug or stereoisomer thereof, wherein:

one of L^1 or L^2 is –O(C=O)-, -(C=O)O-, -C(=O)-, -O-, -S(O)$_x$-, -S-S-, -C(=O)S-, SC(=O)-, -NR^{a}C(=O)-, -C(=O)NR^a-, NR^{a}C(=O)NR^a-, -OC(=O)NR^a- or -NR^{a}C(=O)O-, and the other of L^1 or L^2 is –O(C=O)-, -(C=O)O-, -C(=O)-, -O-, -S(O)$_x$-, -S-S-, -C(=O)S-, SC(=O)-, -NR^{a}C(=O)-, -C(=O)NR^a-, NR^{a}C(=O)NR^a-, -OC(=O)NR^a- or -NR^{a}C(=O)O- or a direct bond;

G^1 and G^2 are each independently unsubstituted C_1-C_{12} alkylene or C_1-C_{12} alkenylene;

G^3 is C_1-C_{24} alkylene, C_1-C_{24} alkenylene, C_3-C_8 cycloalkylene, C_3-C_8 cycloalkenylene;

R^a is H or C_1-C_{12} alkyl;

R^1 and R^2 are each independently C_6-C_{24} alkyl or C_6-C_{24} alkenyl;

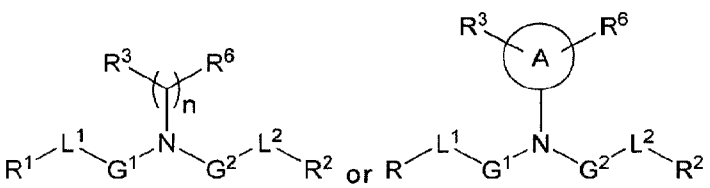
R^3 is H, OR^5, CN, -C(=O)OR^4, -OC(=O)R^4 or –NR^{5}C(=O)R^4;

R^4 is C_1-C_{12} alkyl;

R^5 is H or C_1-C_6 alkyl; and

x is 0, 1 or 2.

In some of the foregoing embodiments of Formula (III), the lipid has one of the following structures (IIIA) or (IIIB):



(IIIA) (IIIB)

wherein:

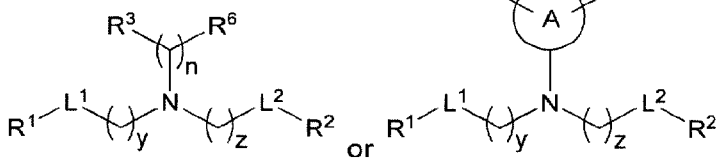
A is a 3 to 8-membered cycloalkyl or cycloalkylene ring;

R^6 is, at each occurrence, independently H, OH or C_1-C_{24} alkyl;

n is an integer ranging from 1 to 15.

In some of the foregoing embodiments of Formula (III), the lipid has structure (IIIA), and in other embodiments, the lipid has structure (IIIB).

In other embodiments of Formula (III), the lipid has one of the following structures (IIIC) or (IIID):

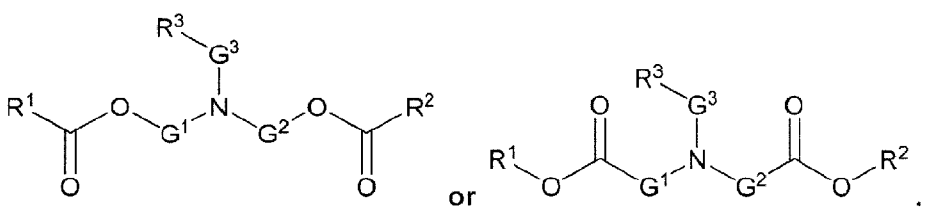


(IIIC) (IIID)

wherein y and z are each independently integers ranging from 1 to 12.

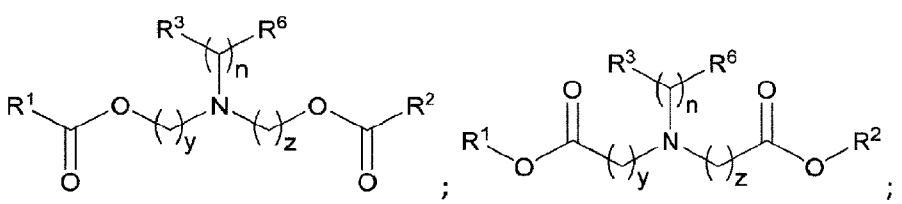
In any of the foregoing embodiments of Formula (III), one of L^1 or L^2 is -O(C=O)-. For example, in some embodiments each of L^1 and L^2 are -O(C=O)-. In some different embodiments of any of the foregoing, L^1 and L^2 are each independently -(C=O)O- or -O(C=O)-. For example, in some embodiments each of L^1 and L^2 is -(C=O)O-.

In some different embodiments of Formula (III), the lipid has one of the following structures (IIIE) or (IIIF):

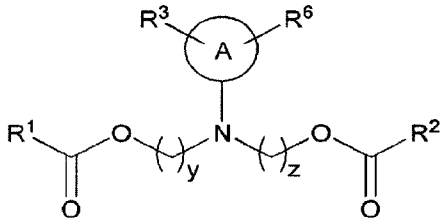


(IIIE) (IIIF)

In some of the foregoing embodiments of Formula (III), the lipid has one of the following structures (IIIG), (IIIH), (IIII), or (IIIJ):

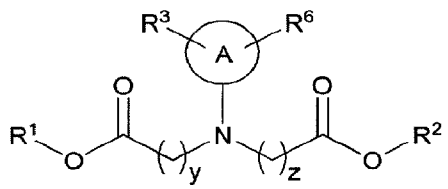


(IIIG) (IIIH)



(IIII)

or



(IIIJ)

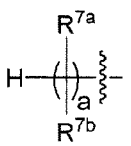
In some of the foregoing embodiments of Formula (III), n is an integer ranging from 2 to 12, for example from 2 to 8 or from 2 to 4. For example, in some embodiments, n is 3, 4, 5 or 6. In some embodiments, n is 3. In some embodiments, n is 4. In some embodiments, n is 5. In some embodiments, n is 6.

In some other of the foregoing embodiments of Formula (III), y and z are each independently an integer ranging from 2 to 10. For example, in some embodiments, y and z are each independently an integer ranging from 4 to 9 or from 4 to 6.

In some of the foregoing embodiments of Formula (III), R^6 is H. In other of the foregoing embodiments, R^6 is C_1-C_{24} alkyl. In other embodiments, R^6 is OH.

In some embodiments of Formula (III), G^3 is unsubstituted. In other embodiments, G3 is substituted. In various different embodiments, G^3 is linear C_1-C_{24} alkylene or linear C_1-C_{24} alkenylene.

In some other foregoing embodiments of Formula (III), R^1 or R^2, or both, is C_6-C_{24} alkenyl. For example, in some embodiments, R^1 and R^2 each, independently have the following structure:



wherein:

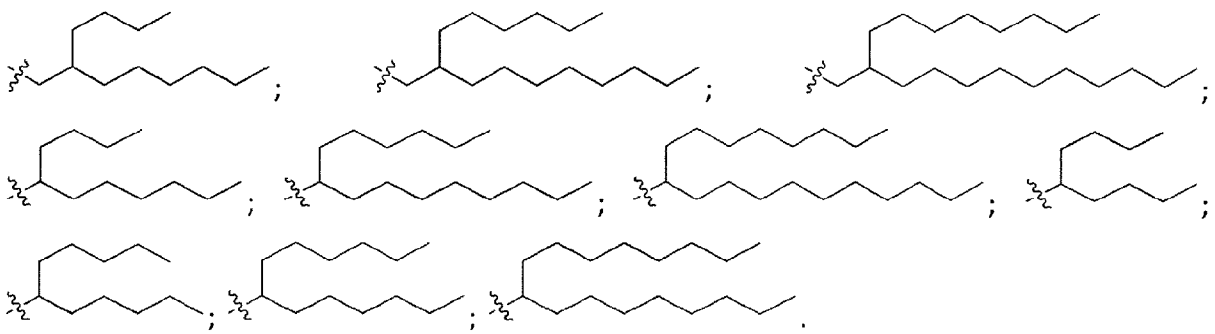
R^{7a} and R^{7b} are, at each occurrence, independently H or C_1-C_{12} alkyl; and

a is an integer from 2 to 12,

wherein R^{7a}, R^{7b} and a are each selected such that R^1 and R^2 each independently comprise from 6 to 20 carbon atoms. For example, in some embodiments a is an integer ranging from 5 to 9 or from 8 to 12.

In some of the foregoing embodiments of Formula (III), at least one occurrence of R^{7a} is H. For example, in some embodiments, R^{7a} is H at each occurrence. In other different embodiments of the foregoing, at least one occurrence of R^{7b} is C_1-C_8 alkyl. For example, in some embodiments, C_1-C_8 alkyl is methyl, ethyl, n-propyl, iso-propyl, n-butyl, iso-butyl, tert-butyl, n-hexyl or n-octyl.

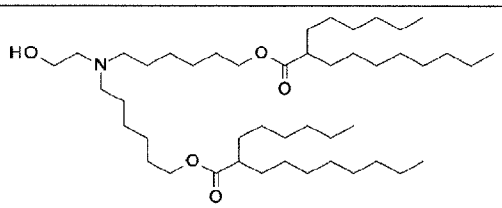
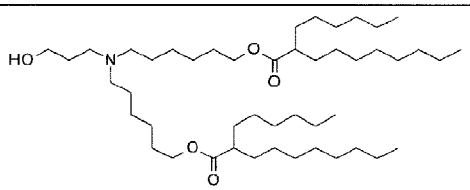
In different embodiments of Formula (III), R^1 or R^2, or both, has one of the following structures:

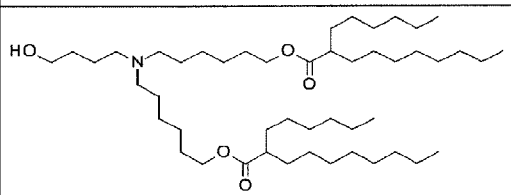
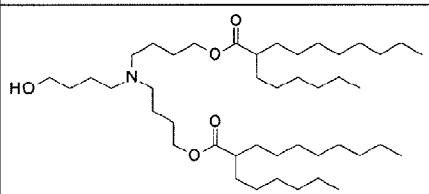
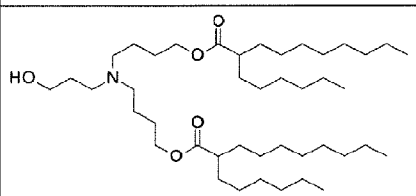
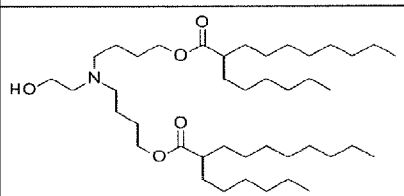
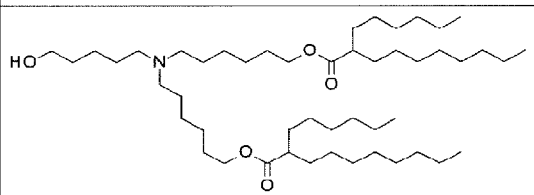
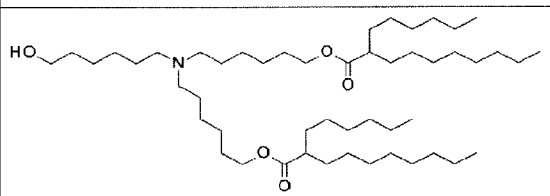
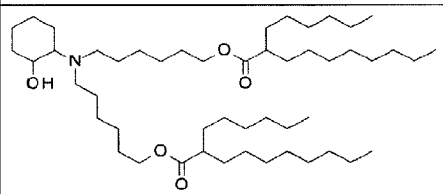


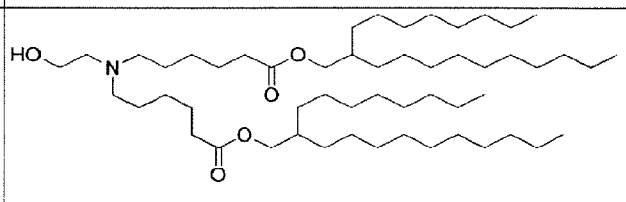
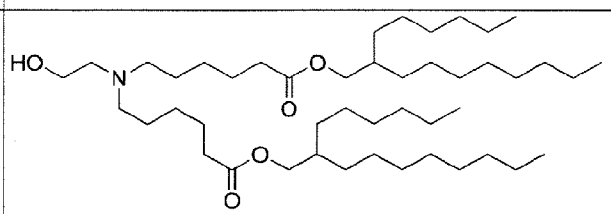
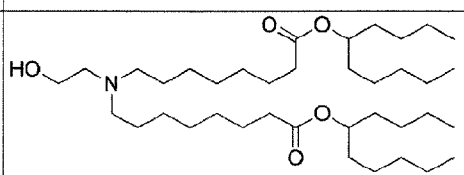
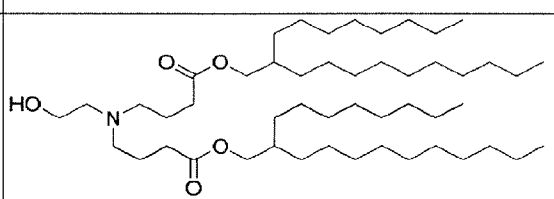
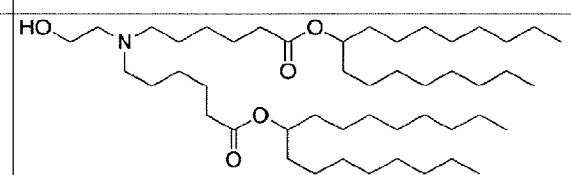
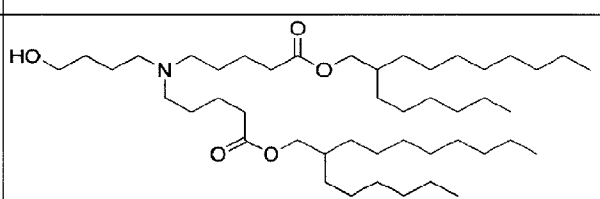
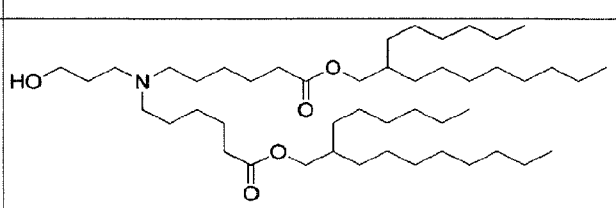
In some of the foregoing embodiments of Formula (III), R^3 is OH, CN, -C(=O)OR^4, -OC(=O)R^4 or –NHC(=O)R^4. In some embodiments, R^4 is methyl or ethyl.

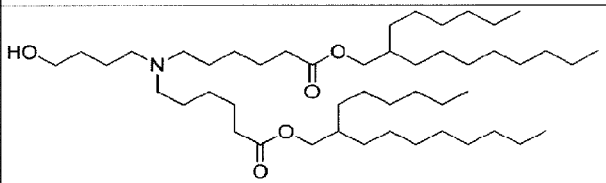
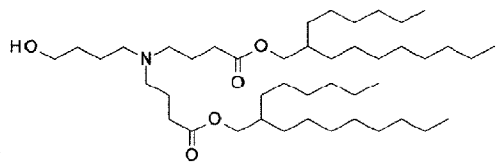
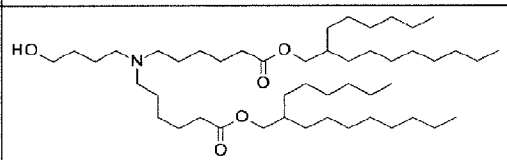
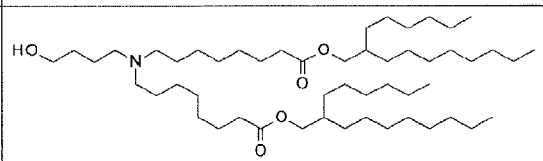
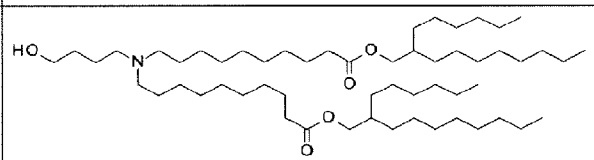
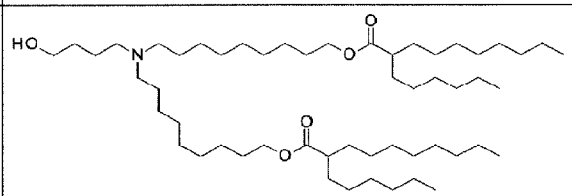
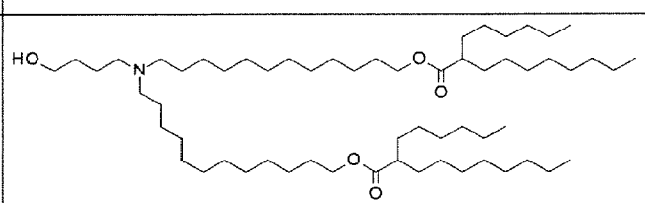
In various different embodiments, the cationic lipid of Formula (III) has one of the structures set forth in the table below.

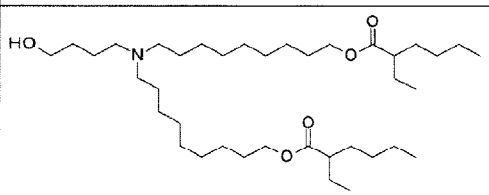
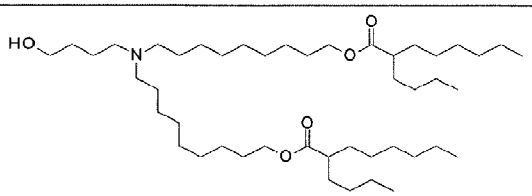
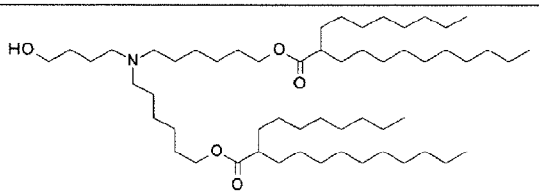
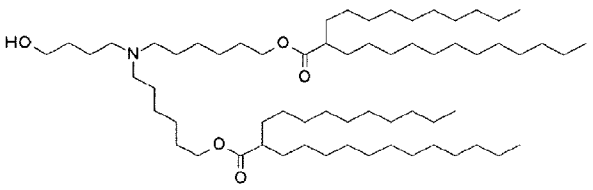
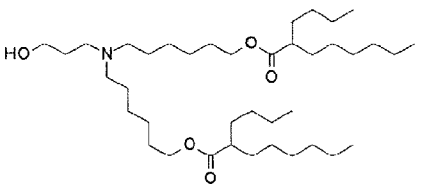
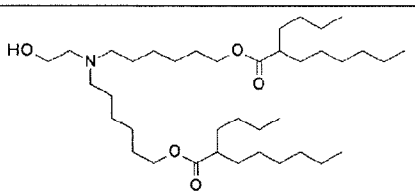
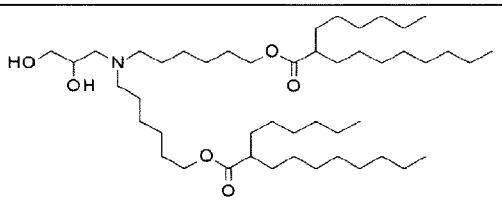
Table 2: Representative Compounds of Formula (III).

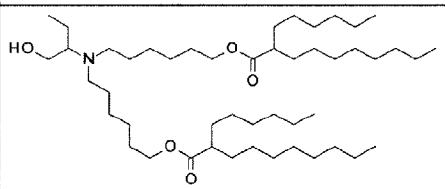
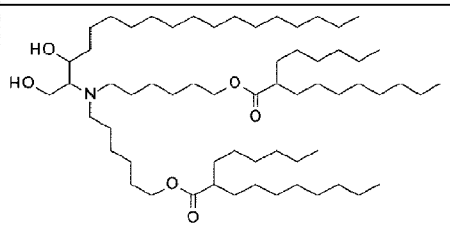
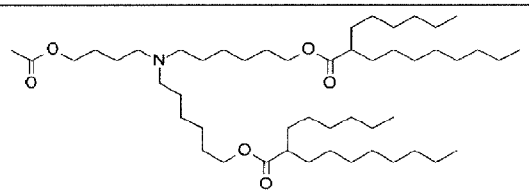
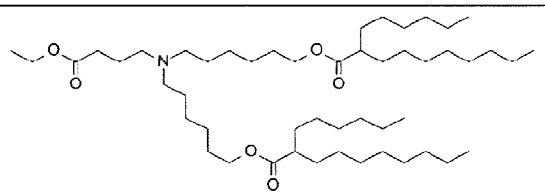
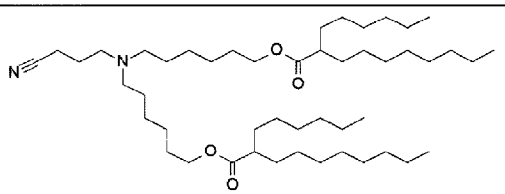
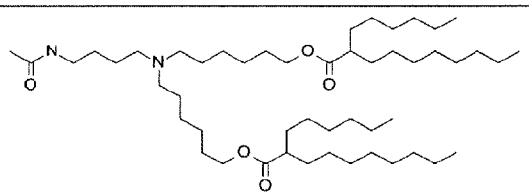
No.	Structure
III-1	
III-2	

No.	Structure
III-3	
III-4	
III-5	
III-6	
III-7	
III-8	
III-9	

No.	Structure
III-10	
III-11	
III-12	
III-13	
III-14	
III-15	
III-16	

No.	Structure
III-17	
III-18	
III-19	
III-20	
III-21	
III-22	
III-23	

No.	Structure
III-24	
III-25	
III-26	
III-27	
III-28	
III-29	
III-30	

No.	Structure
III-31	
III-32	
III-33	
III-34	
III-35	
III-36	

In some embodiments, the LNP comprises a lipid of Formula (III), RNA, a neutral lipid, a steroid and a pegylated lipid. In some embodiments, the lipid of Formula (III) is compound III-3. In

some embodiments, the neutral lipid is DSPC. In some embodiments, the steroid is cholesterol. In some embodiments, the pegylated lipid is ALC-0159.

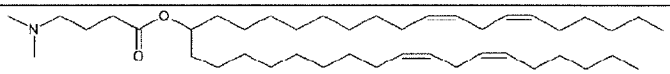
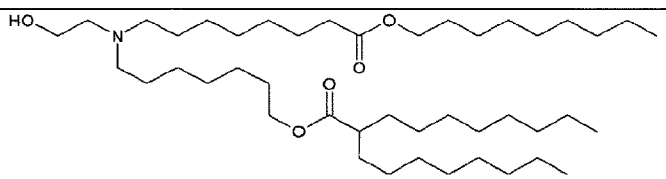
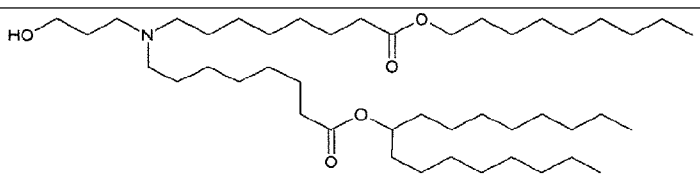
In some embodiments, the cationic lipid is present in the LNP in an amount from about 40 to about 50 mole percent. In one embodiment, the neutral lipid is present in the LNP in an amount from about 5 to about 15 mole percent. In one embodiment, the steroid is present in the LNP in an amount from about 35 to about 45 mole percent. In one embodiment, the pegylated lipid is present in the LNP in an amount from about 1 to about 10 mole percent.

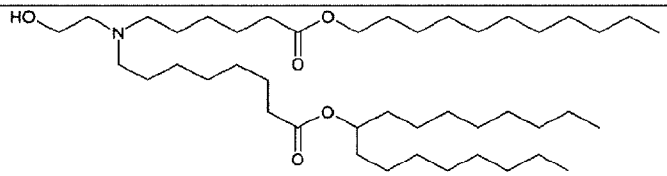
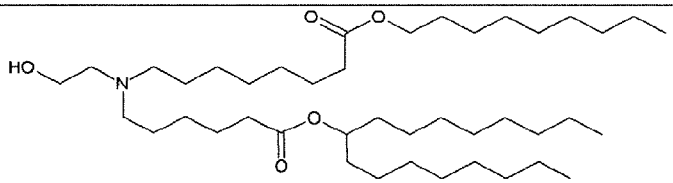
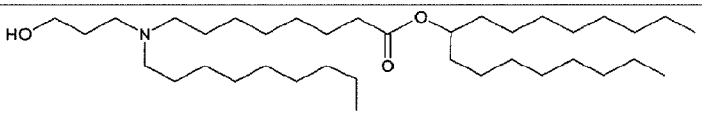
In some embodiments, the LNP comprises compound III-3 in an amount from about 40 to about 50 mole percent, DSPC in an amount from about 5 to about 15 mole percent, cholesterol in an amount from about 35 to about 45 mole percent, and ALC-0159 in an amount from about 1 to about 10 mole percent.

In some embodiments, the LNP comprises compound III-3 in an amount of about 47.5 mole percent, DSPC in an amount of about 10 mole percent, cholesterol in an amount of about 40.7 mole percent, and ALC-0159 in an amount of about 1.8 mole percent.

In various different embodiments, the cationic lipid has one of the structures set forth in the table below.

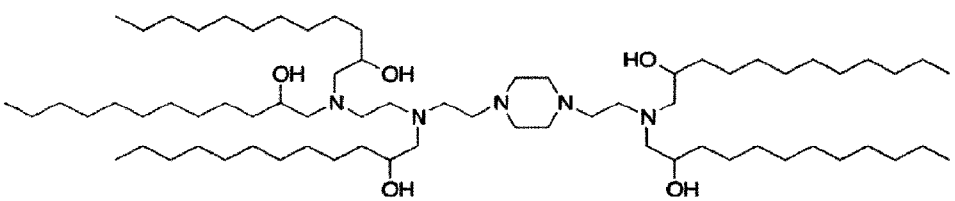
Table 3: Representative cationic lipids.

No.	Structure
A	
B	
C	

No.	Structure
D	
E	
F	

In some embodiments, the LNP comprises a cationic lipid shown in the above table, e.g., a cationic lipid of Formula (B) or Formula (D), in particular a cationic lipid of Formula (D), RNA, a neutral lipid, a steroid and a pegylated lipid. In some embodiments, the neutral lipid is DSPC. In some embodiments, the steroid is cholesterol. In some embodiments, the pegylated lipid is DMG-PEG 2000.

In one embodiment, the LNP comprises a cationic lipid that is an ionizable lipid-like material (lipidoid). In one embodiment, the cationic lipid has the following structure:



The N/P value is preferably at least about 4. In some embodiments, the N/P value ranges from 4 to 20, 4 to 12, 4 to 10, 4 to 8, or 5 to 7. In one embodiment, the N/P value is about 6.

LNP described herein may have an average diameter that in one embodiment ranges from about 30 nm to about 200 nm, or from about 60 nm to about 120 nm.

RNA Targeting

Some aspects of the disclosure involve the targeted delivery of the RNA disclosed herein (e.g., RNA encoding vaccine antigens and/or immunostimulants).

In one embodiment, the disclosure involves targeting lung. Targeting lung is in particular preferred if the RNA administered is RNA encoding vaccine antigen. RNA may be delivered to lung, for example, by administering the RNA which may be formulated as particles as described herein, e.g., lipid particles, by inhalation.

In one embodiment, the disclosure involves targeting the lymphatic system, in particular secondary lymphoid organs, more specifically spleen. Targeting the lymphatic system, in particular secondary lymphoid organs, more specifically spleen is in particular preferred if the RNA administered is RNA encoding vaccine antigen.

In one embodiment, the target cell is a spleen cell. In one embodiment, the target cell is an antigen presenting cell such as a professional antigen presenting cell in the spleen. In one embodiment, the target cell is a dendritic cell in the spleen.

The "lymphatic system" is part of the circulatory system and an important part of the immune system, comprising a network of lymphatic vessels that carry lymph. The lymphatic system consists of lymphatic organs, a conducting network of lymphatic vessels, and the circulating lymph. The primary or central lymphoid organs generate lymphocytes from immature progenitor cells. The thymus and the bone marrow constitute the primary lymphoid organs. Secondary or peripheral lymphoid organs, which include lymph nodes and the spleen, maintain mature naïve lymphocytes and initiate an adaptive immune response.

RNA may be delivered to spleen by so-called lipoplex formulations, in which the RNA is bound to liposomes comprising a cationic lipid and optionally an additional or helper lipid to form injectable nanoparticle formulations. The liposomes may be obtained by injecting a solution of the lipids in ethanol into water or a suitable aqueous phase. RNA lipoplex particles may be prepared by mixing the liposomes with RNA. Spleen targeting RNA lipoplex particles are described in WO 2013/143683, herein incorporated by reference. It has been found that RNA lipoplex particles having a net negative charge may be used to preferentially target spleen tissue or spleen cells such as antigen-presenting cells, in particular dendritic cells. Accordingly, following administration of the RNA lipoplex particles, RNA accumulation and/or RNA expression in the spleen occurs. Thus, RNA lipoplex particles of the disclosure may be used for expressing RNA in the spleen. In an embodiment, after administration of the RNA lipoplex particles, no or essentially no RNA accumulation and/or RNA expression in the lung and/or

CA 03215771 2023-10-17

liver occurs. In one embodiment, after administration of the RNA lipoplex particles, RNA accumulation and/or RNA expression in antigen presenting cells, such as professional antigen presenting cells in the spleen occurs. Thus, RNA lipoplex particles of the disclosure may be used for expressing RNA in such antigen presenting cells. In one embodiment, the antigen presenting cells are dendritic cells and/or macrophages.

The electric charge of the RNA lipoplex particles of the present disclosure is the sum of the electric charges present in the at least one cationic lipid and the electric charges present in the RNA. The charge ratio is the ratio of the positive charges present in the at least one cationic lipid to the negative charges present in the RNA. The charge ratio of the positive charges present in the at least one cationic lipid to the negative charges present in the RNA is calculated by the following equation: charge ratio=[(cationic lipid concentration (mol)) * (the total number of positive charges in the cationic lipid)] / [(RNA concentration (mol)) * (the total number of negative charges in RNA)].

The spleen targeting RNA lipoplex particles described herein at physiological pH preferably have a net negative charge such as a charge ratio of positive charges to negative charges from about 1.9:2 to about 1:2, or about 1.6:2 to about 1:2, or about 1.6:2 to about 1.1:2. In specific embodiments, the charge ratio of positive charges to negative charges in the RNA lipoplex particles at physiological pH is about 1.9:2.0, about 1.8:2.0, about 1.7:2.0, about 1.6:2.0, about 1.5:2.0, about 1.4:2.0, about 1.3:2.0, about 1.2:2.0, about 1.1:2.0, or about 1:2.0.

Immunostimulants may be provided to a subject by administering to the subject RNA encoding an immunostimulant in a formulation for preferential delivery of RNA to liver or liver tissue. The delivery of RNA to such target organ or tissue is preferred, in particular, if it is desired to express large amounts of the immunostimulant and/or if systemic presence of the immunostimulant, in particular in significant amounts, is desired or required.

RNA delivery systems have an inherent preference to the liver. This pertains to lipid-based particles, cationic and neutral nanoparticles, in particular lipid nanoparticles such as liposomes, nanomicelles and lipophilic ligands in bioconjugates. Liver accumulation is caused by the discontinuous nature of the hepatic vasculature or the lipid metabolism (liposomes and lipid or cholesterol conjugates).

For *in vivo* delivery of RNA to the liver, a drug delivery system may be used to transport the RNA into the liver by preventing its degradation. For example, polyplex nanomicelles consisting of a poly(ethylene glycol) (PEG)-coated surface and an mRNA-containing core is a useful system because the nanomicelles provide excellent *in vivo* stability of the RNA, under physiological conditions. Furthermore, the stealth property provided by the polyplex nanomicelle surface, composed of dense PEG palisades, effectively evades host immune defenses.

Examples of suitable immunostimulants for targeting liver are cytokines involved in T cell proliferation and/or maintenance. Examples of suitable cytokines include IL2 or IL7, fragments and variants thereof, and fusion proteins of these cytokines, fragments and variants, such as extended-PK cytokines.

In another embodiment, RNA encoding an immunostimulant may be administered in a formulation for preferential delivery of RNA to the lymphatic system, in particular secondary lymphoid organs, more specifically spleen. The delivery of an immunostimulant to such target tissue is preferred, in particular, if presence of the immunostimulant in this organ or tissue is desired (e.g., for inducing an immune response, in particular in case immunostimulants such as cytokines are required during T-cell priming or for activation of resident immune cells), while it is not desired that the immunostimulant is present systemically, in particular in significant amounts (e.g., because the immunostimulant has systemic toxicity).

Examples of suitable immunostimulants are cytokines involved in T cell priming. Examples of suitable cytokines include IL12, IL15, IFN-α, or IFN-β, fragments and variants thereof, and fusion proteins of these cytokines, fragments and variants, such as extended-PK cytokines.

Immunostimulants

In one embodiment, the RNA encoding vaccine antigen may be non-immunogenic. In this and other embodiments, the RNA encoding vaccine antigen may be co-administered with an immunostimulant or RNA encoding an immunostimulant. The methods and agents described herein are particularly effective if the immunostimulant is attached to a pharmacokinetic modifying group (hereafter referred to as "extended-pharmacokinetic (PK)" immunostimulant). The methods and agents described herein are particularly effective if the

CA 03215771 2023-10-17

immunostimulant is administered in the form of RNA encoding an immunostimulant. In one embodiment, said RNA is targeted to the liver for systemic availability. Liver cells can be efficiently transfected and are able to produce large amounts of protein.

An "immunostimulant" is any substance that stimulates the immune system by inducing activation or increasing activity of any of the immune system's components, in particular immune effector cells. The immunostimulant may be pro-inflammatory.

According to one aspect, the immunostimulant is a cytokine or a variant thereof. Examples of cytokines include interferons, such as interferon-alpha (IFN-α) or interferon-gamma (IFN-γ), interleukins, such as IL2, IL7, IL12, IL15 and IL23, colony stimulating factors, such as M-CSF and GM-CSF, and tumor necrosis factor. According to another aspect, the immunostimulant includes an adjuvant-type immunostimulatory agent such as APC Toll-like Receptor agonists or costimulatory/cell adhesion membrane proteins. Examples of Toll-like Receptor agonists include costimulatory/adhesion proteins such as CD80, CD86, and ICAM-1.

Cytokines are a category of small proteins (~5–20 kDa) that are important in cell signaling. Their release has an effect on the behavior of cells around them. Cytokines are involved in autocrine signaling, paracrine signaling and endocrine signaling as immunomodulating agents. Cytokines include chemokines, interferons, interleukins, lymphokines, and tumour necrosis factors but generally not hormones or growth factors (despite some overlap in the terminology). Cytokines are produced by a broad range of cells, including immune cells like macrophages, B lymphocytes, T lymphocytes and mast cells, as well as endothelial cells, fibroblasts, and various stromal cells. A given cytokine may be produced by more than one type of cell. Cytokines act through receptors, and are especially important in the immune system; cytokines modulate the balance between humoral and cell-based immune responses, and they regulate the maturation, growth, and responsiveness of particular cell populations. Some cytokines enhance or inhibit the action of other cytokines in complex ways.

According to the disclosure, a cytokine may be a naturally occurring cytokine or a functional fragment or variant thereof. A cytokine may be human cytokine and may be derived from any vertebrate, especially any mammal. One particularly preferred cytokine is interferon-α.

Interferons

Interferons (IFNs) are a group of signaling proteins made and released by host cells in response to the presence of several pathogens, such as viruses, bacteria, parasites, and also tumor cells. In a typical scenario, a virus-infected cell will release interferons causing nearby cells to heighten their anti-viral defenses.

Based on the type of receptor through which they signal, interferons are typically divided among three classes: type I interferon, type II interferon, and type III interferon.

All type I interferons bind to a specific cell surface receptor complex known as the IFN-α/β receptor (IFNAR) that consists of IFNAR1 and IFNAR2 chains.

The type I interferons present in humans are IFNα, IFNβ, IFNε, IFNκ and IFNω. In general, type I interferons are produced when the body recognizes a virus that has invaded it. They are produced by fibroblasts and monocytes. Once released, type I interferons bind to specific receptors on target cells, which leads to expression of proteins that will prevent the virus from producing and replicating its RNA and DNA.

The IFNα proteins are produced mainly by plasmacytoid dendritic cells (pDCs). They are mainly involved in innate immunity against viral infection. The genes responsible for their synthesis come in 13 subtypes that are called IFNA1, IFNA2, IFNA4, IFNA5, IFNA6, IFNA7, IFNA8, IFNA10, IFNA13, IFNA14, IFNA16, IFNA17, IFNA21. These genes are found together in a cluster on chromosome 9.

The IFNβ proteins are produced in large quantities by fibroblasts. They have antiviral activity that is involved mainly in innate immune response. Two types of IFNβ have been described, IFNβ1 and IFNβ3. The natural and recombinant forms of IFNβ1 have antiviral, antibacterial, and anticancer properties.

Type II interferon (IFNγ in humans) is also known as immune interferon and is activated by IL12. Furthermore, type II interferons are released by cytotoxic T cells and T helper cells.

Type III interferons signal through a receptor complex consisting of IL10R2 (also called CRF2-4) and IFNLR1 (also called CRF2-12). Although discovered more recently than type I and type II IFNs, recent information demonstrates the importance of type III IFNs in some types of virus or fungal infections.

CA 03215771 2023-10-17

In general, type I and II interferons are responsible for regulating and activating the immune response.

According to the disclosure, a type I interferon is preferably IFNα or IFNβ, more preferably IFNα.

According to the disclosure, an interferon may be a naturally occurring interferon or a functional fragment or variant thereof. An interferon may be human interferon and may be derived from any vertebrate, especially any mammal.

Interleukins

Interleukins (ILs) are a group of cytokines (secreted proteins and signal molecules) that can be divided into four major groups based on distinguishing structural features. However, their amino acid sequence similarity is rather weak (typically 15–25% identity). The human genome encodes more than 50 interleukins and related proteins.

According to the disclosure, an interleukin may be a naturally occurring interleukin or a functional fragment or variant thereof. An interleukin may be human interleukin and may be derived from any vertebrate, especially any mammal.

Extended-PK group

Immunostimulant polypeptides described herein can be prepared as fusion or chimeric polypeptides that include an immunostimulant portion and a heterologous polypeptide (i.e., a polypeptide that is not an immunostimulant). The immunostimulant may be fused to an extended-PK group, which increases circulation half-life. Non-limiting examples of extended-PK groups are described infra. It should be understood that other PK groups that increase the circulation half-life of immunostimulants such as cytokines, or variants thereof, are also applicable to the present disclosure. In certain embodiments, the extended-PK group is a serum albumin domain (e.g., mouse serum albumin, human serum albumin).

As used herein, the term "PK" is an acronym for "pharmacokinetic" and encompasses properties of a compound including, by way of example, absorption, distribution, metabolism, and elimination by a subject. As used herein, an "extended-PK group" refers to a protein, peptide, or moiety that increases the circulation half-life of a biologically active molecule when

CA 03215771 2023-10-17

fused to or administered together with the biologically active molecule. Examples of an extended-PK group include serum albumin (e.g., HSA), Immunoglobulin Fc or Fc fragments and variants thereof, transferrin and variants thereof, and human serum albumin (HSA) binders (as disclosed in U.S. Publication Nos. 2005/0287153 and 2007/0003549). Other exemplary extended-PK groups are disclosed in Kontermann, Expert Opin Biol Ther, 2016 Jul;16(7):903-15 which is herein incorporated by reference in its entirety. As used herein, an "extended-PK" immunostimulant refers to an immunostimulant moiety in combination with an extended-PK group. In one embodiment, the extended-PK immunostimulant is a fusion protein in which an immunostimulant moiety is linked or fused to an extended-PK group.

In certain embodiments, the serum half-life of an extended-PK immunostimulant is increased relative to the immunostimulant alone (i.e., the immunostimulant not fused to an extended-PK group). In certain embodiments, the serum half-life of the extended-PK immunostimulant is at least 20, 40, 60, 80, 100, 120, 150, 180, 200, 400, 600, 800, or 1000% longer relative to the serum half-life of the immunostimulant alone. In certain embodiments, the serum half-life of the extended-PK immunostimulant is at least 1.5-fold, 2-fold, 2.5-fold, 3-fold, 3.5 fold, 4-fold, 4.5-fold, 5-fold, 6-fold, 7-fold, 8-fold, 10- fold, 12-fold, 13-fold, 15-fold, 17-fold, 20-fold, 22- fold, 25-fold, 27-fold, 30-fold, 35-fold, 40-fold, or 50-fold greater than the serum half-life of the immunostimulant alone. In certain embodiments, the serum half-life of the extended-PK immunostimulant is at least 10 hours, 15 hours, 20 hours, 25 hours, 30 hours, 35 hours, 40 hours, 50 hours, 60 hours, 70 hours, 80 hours, 90 hours, 100 hours, 110 hours, 120 hours, 130 hours, 135 hours, 140 hours, 150 hours, 160 hours, or 200 hours.

As used herein, "half-life" refers to the time taken for the serum or plasma concentration of a compound such as a peptide or protein to reduce by 50%, *in vivo*, for example due to degradation and/or clearance or sequestration by natural mechanisms. An extended-PK immunostimulant suitable for use herein is stabilized *in vivo* and its half-life increased by, e.g., fusion to serum albumin (e.g., HSA or MSA), which resist degradation and/or clearance or sequestration. The half-life can be determined in any manner known per se, such as by pharmacokinetic analysis. Suitable techniques will be clear to the person skilled in the art, and may for example generally involve the steps of suitably administering a suitable dose of the amino acid sequence or compound to a subject; collecting blood samples or other samples

CA 03215771 2023-10-17

from said subject at regular intervals; determining the level or concentration of the amino acid sequence or compound in said blood sample; and calculating, from (a plot of) the data thus obtained, the time until the level or concentration of the amino acid sequence or compound has been reduced by 50% compared to the initial level upon dosing. Further details are provided in, e.g., standard handbooks, such as Kenneth, A. et al., Chemical Stability of Pharmaceuticals: A Handbook for Pharmacists and in Peters et al., Pharmacokinetic Analysis: A Practical Approach (1996). Reference is also made to Gibaldi, M. et al., Pharmacokinetics, 2nd Rev. Edition, Marcel Dekker (1982).

In certain embodiments, the extended-PK group includes serum albumin, or fragments thereof or variants of the serum albumin or fragments thereof (all of which for the purpose of the present disclosure are comprised by the term "albumin"). Polypeptides described herein may be fused to albumin (or a fragment or variant thereof) to form albumin fusion proteins. Such albumin fusion proteins are described in U.S. Publication No. 20070048282.

As used herein, "albumin fusion protein" refers to a protein formed by the fusion of at least one molecule of albumin (or a fragment or variant thereof) to at least one molecule of a protein such as a therapeutic protein, in particular an immunostimulant. The albumin fusion protein may be generated by translation of a nucleic acid in which a polynucleotide encoding a therapeutic protein is joined in-frame with a polynucleotide encoding an albumin. The therapeutic protein and albumin, once part of the albumin fusion protein, may each be referred to as a "portion", "region" or "moiety" of the albumin fusion protein (e.g., a "therapeutic protein portion" or an "albumin protein portion"). In a highly preferred embodiment, an albumin fusion protein comprises at least one molecule of a therapeutic protein (including, but not limited to a mature form of the therapeutic protein) and at least one molecule of albumin (including but not limited to a mature form of albumin). In one embodiment, an albumin fusion protein is processed by a host cell such as a cell of the target organ for administered RNA, e.g. a liver cell, and secreted into the circulation. Processing of the nascent albumin fusion protein that occurs in the secretory pathways of the host cell used for expression of the RNA may include, but is not limited to signal peptide cleavage; formation of disulfide bonds; proper folding; addition and processing of carbohydrates (such as for example, N- and O-linked glycosylation); specific proteolytic cleavages; and/or assembly into

CA 03215771 2023-10-17

multimeric proteins. An albumin fusion protein is preferably encoded by RNA in a non-processed form which in particular has a signal peptide at its N-terminus and following secretion by a cell is preferably present in the processed form wherein in particular the signal peptide has been cleaved off. In a most preferred embodiment, the "processed form of an albumin fusion protein" refers to an albumin fusion protein product which has undergone N-terminal signal peptide cleavage, herein also referred to as a "mature albumin fusion protein". In preferred embodiments, albumin fusion proteins comprising a therapeutic protein have a higher plasma stability compared to the plasma stability of the same therapeutic protein when not fused to albumin. Plasma stability typically refers to the time period between when the therapeutic protein is administered *in vivo* and carried into the bloodstream and when the therapeutic protein is degraded and cleared from the bloodstream, into an organ, such as the kidney or liver, that ultimately clears the therapeutic protein from the body. Plasma stability is calculated in terms of the half-life of the therapeutic protein in the bloodstream. The half-life of the therapeutic protein in the bloodstream can be readily determined by common assays known in the art.

As used herein, "albumin" refers collectively to albumin protein or amino acid sequence, or an albumin fragment or variant, having one or more functional activities (e.g., biological activities) of albumin. In particular, "albumin" refers to human albumin or fragments or variants thereof especially the mature form of human albumin, or albumin from other vertebrates or fragments thereof, or variants of these molecules. The albumin may be derived from any vertebrate, especially any mammal, for example human, cow, sheep, or pig. Non-mammalian albumins include, but are not limited to, hen and salmon. The albumin portion of the albumin fusion protein may be from a different animal than the therapeutic protein portion.

In certain embodiments, the albumin is human serum albumin (HSA), or fragments or variants thereof, such as those disclosed in US 5,876,969, WO 2011/124718, WO 2013/075066, and WO 2011/0514789.

The terms, human serum albumin (HSA) and human albumin (HA) are used interchangeably herein. The terms, "albumin and "serum albumin" are broader, and encompass human serum

CA 03215771 2023-10-17

albumin (and fragments and variants thereof) as well as albumin from other species (and fragments and variants thereof).

As used herein, a fragment of albumin sufficient to prolong the therapeutic activity or plasma stability of the therapeutic protein refers to a fragment of albumin sufficient in length or structure to stabilize or prolong the therapeutic activity or plasma stability of the protein so that the plasma stability of the therapeutic protein portion of the albumin fusion protein is prolonged or extended compared to the plasma stability in the non-fusion state.

The albumin portion of the albumin fusion proteins may comprise the full length of the albumin sequence, or may include one or more fragments thereof that are capable of stabilizing or prolonging the therapeutic activity or plasma stability. Such fragments may be of 10 or more amino acids in length or may include about 15, 20, 25, 30, 50, or more contiguous amino acids from the albumin sequence or may include part or all of specific domains of albumin. For instance, one or more fragments of HSA spanning the first two immunoglobulin-like domains may be used. In a preferred embodiment, the HSA fragment is the mature form of HSA.

Generally speaking, an albumin fragment or variant will be at least 100 amino acids long, preferably at least 150 amino acids long.

According to the disclosure, albumin may be naturally occurring albumin or a fragment or variant thereof. Albumin may be human albumin and may be derived from any vertebrate, especially any mammal.

Preferably, the albumin fusion protein comprises albumin as the N-terminal portion, and a therapeutic protein as the C-terminal portion. Alternatively, an albumin fusion protein comprising albumin as the C-terminal portion, and a therapeutic protein as the N-terminal portion may also be used. In other embodiments, the albumin fusion protein has a therapeutic protein fused to both the N-terminus and the C-terminus of albumin. In a preferred embodiment, the therapeutic proteins fused at the N- and C-termini are the same therapeutic proteins. In another preferred embodiment, the therapeutic proteins fused at the N- and C-termini are different therapeutic proteins. In one embodiment, the different therapeutic proteins are both cytokines.

In one embodiment, the therapeutic protein(s) is (are) joined to the albumin through (a) peptide linker(s). A linker peptide between the fused portions may provide greater physical separation between the moieties and thus maximize the accessibility of the therapeutic protein portion, for instance, for binding to its cognate receptor. The linker peptide may consist of amino acids such that it is flexible or more rigid. The linker sequence may be cleavable by a protease or chemically.

As used herein, the term "Fc region" refers to the portion of a native immunoglobulin formed by the respective Fc domains (or Fc moieties) of its two heavy chains. As used herein, the term "Fc domain" refers to a portion or fragment of a single immunoglobulin (Ig) heavy chain wherein the Fc domain does not comprise an Fv domain. In certain embodiments, an Fc domain begins in the hinge region just upstream of the papain cleavage site and ends at the C-terminus of the antibody. Accordingly, a complete Fc domain comprises at least a hinge domain, a CH2 domain, and a CH3 domain. In certain embodiments, an Fc domain comprises at least one of: a hinge (e.g., upper, middle, and/or lower hinge region) domain, a CH2 domain, a CH3 domain, a CH4 domain, or a variant, portion, or fragment thereof. In certain embodiments, an Fc domain comprises a complete Fc domain (i.e., a hinge domain, a CH2 domain, and a CH3 domain). In certain embodiments, an Fc domain comprises a hinge domain (or portion thereof) fused to a CH3 domain (or portion thereof). In certain embodiments, an Fc domain comprises a CH2 domain (or portion thereof) fused to a CH3 domain (or portion thereof). In certain embodiments, an Fc domain consists of a CH3 domain or portion thereof. In certain embodiments, an Fc domain consists of a hinge domain (or portion thereof) and a CH3 domain (or portion thereof). In certain embodiments, an Fc domain consists of a CH2 domain (or portion thereof) and a CH3 domain. In certain embodiments, an Fc domain consists of a hinge domain (or portion thereof) and a CH2 domain (or portion thereof). In certain embodiments, an Fc domain lacks at least a portion of a CH2 domain (e.g., all or part of a CH2 domain). An Fc domain herein generally refers to a polypeptide comprising all or part of the Fc domain of an immunoglobulin heavy-chain. This includes, but is not limited to, polypeptides comprising the entire CH1, hinge, CH2, and/or CH3 domains as well as fragments of such peptides comprising only, e.g., the hinge, CH2, and CH3 domain. The Fc domain may be derived from an immunoglobulin of any species and/or any subtype, including, but not limited

CA 03215771 2023-10-17

to, a human IgG1, IgG2, IgG3, IgG4, IgD, IgA, IgE, or IgM antibody. The Fc domain encompasses native Fc and Fc variant molecules. As set forth herein, it will be understood by one of ordinary skill in the art that any Fc domain may be modified such that it varies in amino acid sequence from the native Fc domain of a naturally occurring immunoglobulin molecule. In certain embodiments, the Fc domain has reduced effector function (e.g., FcγR binding).

The Fc domains of a polypeptide described herein may be derived from different immunoglobulin molecules. For example, an Fc domain of a polypeptide may comprise a CH2 and/or CH3 domain derived from an IgG1 molecule and a hinge region derived from an IgG3 molecule. In another example, an Fc domain can comprise a chimeric hinge region derived, in part, from an IgG1 molecule and, in part, from an IgG3 molecule. In another example, an Fc domain can comprise a chimeric hinge derived, in part, from an IgG1 molecule and, in part, from an IgG4 molecule.

In certain embodiments, an extended-PK group includes an Fc domain or fragments thereof or variants of the Fc domain or fragments thereof (all of which for the purpose of the present disclosure are comprised by the term "Fc domain"). The Fc domain does not contain a variable region that binds to antigen. Fc domains suitable for use in the present disclosure may be obtained from a number of different sources. In certain embodiments, an Fc domain is derived from a human immunoglobulin. In certain embodiments, the Fc domain is from a human IgG1 constant region. It is understood, however, that the Fc domain may be derived from an immunoglobulin of another mammalian species, including for example, a rodent (e.g. a mouse, rat, rabbit, guinea pig) or non- human primate (e.g. chimpanzee, macaque) species.

Moreover, the Fc domain (or a fragment or variant thereof) may be derived from any immunoglobulin class, including IgM, IgG, IgD, IgA, and IgE, and any immunoglobulin isotype, including IgG1, IgG2, IgG3, and IgG4.

A variety of Fc domain gene sequences (e.g., mouse and human constant region gene sequences) are available in the form of publicly accessible deposits. Constant region domains comprising an Fc domain sequence can be selected lacking a particular effector function and/or with a particular modification to reduce immunogenicity. Many sequences of antibodies and antibody-encoding genes have been published and suitable Fc domain

CA 03215771 2023-10-17

sequences (e.g. hinge, CH2, and/or CH3 sequences, or fragments or variants thereof) can be derived from these sequences using art recognized techniques.

In certain embodiments, the extended-PK group is a serum albumin binding protein such as those described in US2005/0287153, US2007/0003549, US2007/0178082, US2007/0269422, US2010/0113339, WO2009/083804, and WO2009/133208, which are herein incorporated by reference in their entirety. In certain embodiments, the extended-PK group is transferrin, as disclosed in US 7,176,278 and US 8,158,579, which are herein incorporated by reference in their entirety. In certain embodiments, the extended-PK group is a serum immunoglobulin binding protein such as those disclosed in US2007/0178082, US2014/0220017, and US2017/0145062, which are herein incorporated by reference in their entirety. In certain embodiments, the extended-PK group is a fibronectin (Fn)-based scaffold domain protein that binds to serum albumin, such as those disclosed in US2012/0094909, which is herein incorporated by reference in its entirety. Methods of making fibronectin-based scaffold domain proteins are also disclosed in US2012/0094909. A non-limiting example of a Fn3-based extended-PK group is Fn3(HSA), i.e., a Fn3 protein that binds to human serum albumin.

In certain aspects, the extended-PK immunostimulant, suitable for use according to the disclosure, can employ one or more peptide linkers. As used herein, the term "peptide linker" refers to a peptide or polypeptide sequence which connects two or more domains (e.g., the extended-PK moiety and an immunostimulant moiety) in a linear amino acid sequence of a polypeptide chain. For example, peptide linkers may be used to connect an immunostimulant moiety to a HSA domain.

Linkers suitable for fusing the extended-PK group to e.g. an immunostimulant are well known in the art. Exemplary linkers include glycine-serine-polypeptide linkers, glycine-proline-polypeptide linkers, and proline-alanine polypeptide linkers. In certain embodiments, the linker is a glycine-serine-polypeptide linker, i.e., a peptide that consists of glycine and serine residues.

In addition to, or in place of, the heterologous polypeptides described above, an immunostimulant polypeptide described herein can contain sequences encoding a "marker" or "reporter". Examples of marker or reporter genes include β-lactamase, chloramphenicol acetyltransferase (CAT), adenosine deaminase (ADA), aminoglycoside phosphotransferase,

CA 03215771 2023-10-17

dihydrofolate reductase (DHFR), hygromycin-B-hosphotransferase (HPH), thymidine kinase (TK), β-galactosidase, and xanthine guanine phosphoribosyltransferase (XGPRT).

Pharmaceutical compositions

The agents described herein may be administered in pharmaceutical compositions or medicaments and may be administered in the form of any suitable pharmaceutical composition.

In one embodiment, the pharmaceutical composition described herein is an immunogenic composition for inducing an immune response against a virus, e.g., coronavirus in a subject. For example, in one embodiment, the immunogenic composition is a vaccine.

In one embodiment of all aspects of the invention, the components described herein such as RNA encoding a vaccine antigen may be administered in a pharmaceutical composition which may comprise a pharmaceutically acceptable carrier and may optionally comprise one or more adjuvants, stabilizers etc. In one embodiment, the pharmaceutical composition is for therapeutic or prophylactic treatments, e.g., for use in treating or preventing virus infection, e.g., coronavirus infection.

The term "pharmaceutical composition" relates to a formulation comprising a therapeutically effective agent, preferably together with pharmaceutically acceptable carriers, diluents and/or excipients. Said pharmaceutical composition is useful for treating, preventing, or reducing the severity of a disease or disorder by administration of said pharmaceutical composition to a subject. A pharmaceutical composition is also known in the art as a pharmaceutical formulation.

The pharmaceutical compositions of the present disclosure may comprise one or more adjuvants or may be administered with one or more adjuvants. The term "adjuvant" relates to a compound which prolongs, enhances or accelerates an immune response. Adjuvants comprise a heterogeneous group of compounds such as oil emulsions (e.g., Freund's adjuvants), mineral compounds (such as alum), bacterial products (such as Bordetella pertussis toxin), or immune-stimulating complexes. Examples of adjuvants include, without limitation, LPS, GP96, CpG oligodeoxynucleotides, growth factors, and cytokines, such as monokines, lymphokines, interleukins, chemokines. The cytokines may be IL1, IL2, IL3, IL4, IL5,

CA 03215771 2023-10-17

IL6, IL7, IL8, IL9, IL10, IL12, IFNα, IFNγ, GM-CSF, LT-a. Further known adjuvants are aluminium hydroxide, Freund's adjuvant or oil such as Montanide® ISA51. Other suitable adjuvants for use in the present disclosure include lipopeptides, such as Pam3Cys.

The pharmaceutical compositions according to the present disclosure are generally applied in a "pharmaceutically effective amount" and in "a pharmaceutically acceptable preparation".

The term "pharmaceutically acceptable" refers to the non-toxicity of a material which does not interact with the action of the active component of the pharmaceutical composition.

The term "pharmaceutically effective amount" or "therapeutically effective amount" refers to the amount which achieves a desired reaction or a desired effect alone or together with further doses. In the case of the treatment of a particular disease, the desired reaction preferably relates to inhibition of the course of the disease. This comprises slowing down the progress of the disease and, in particular, interrupting or reversing the progress of the disease. The desired reaction in a treatment of a disease may also be delay of the onset or a prevention of the onset of said disease or said condition. An effective amount of the compositions described herein will depend on the condition to be treated, the severeness of the disease, the individual parameters of the patient, including age, physiological condition, size and weight, the duration of treatment, the type of an accompanying therapy (if present), the specific route of administration and similar factors. Accordingly, the doses administered of the compositions described herein may depend on various of such parameters. In the case that a reaction in a patient is insufficient with an initial dose, higher doses (or effectively higher doses achieved by a different, more localized route of administration) may be used.

The pharmaceutical compositions of the present disclosure may contain salts, buffers, preservatives, and optionally other therapeutic agents. In one embodiment, the pharmaceutical compositions of the present disclosure comprise one or more pharmaceutically acceptable carriers, diluents and/or excipients.

Suitable preservatives for use in the pharmaceutical compositions of the present disclosure include, without limitation, benzalkonium chloride, chlorobutanol, paraben and thimerosal.

The term "excipient" as used herein refers to a substance which may be present in a pharmaceutical composition of the present disclosure but is not an active ingredient. Examples of excipients, include without limitation, carriers, binders, diluents, lubricants,

CA 03215771 2023-10-17

thickeners, surface active agents, preservatives, stabilizers, emulsifiers, buffers, flavoring agents, or colorants.

The term "diluent" relates a diluting and/or thinning agent. Moreover, the term "diluent" includes any one or more of fluid, liquid or solid suspension and/or mixing media. Examples of suitable diluents include ethanol, glycerol and water.

The term "carrier" refers to a component which may be natural, synthetic, organic, inorganic in which the active component is combined in order to facilitate, enhance or enable administration of the pharmaceutical composition. A carrier as used herein may be one or more compatible solid or liquid fillers, diluents or encapsulating substances, which are suitable for administration to subject. Suitable carrier include, without limitation, sterile water, Ringer, Ringer lactate, sterile sodium chloride solution, isotonic saline, polyalkylene glycols, hydrogenated naphthalenes and, in particular, biocompatible lactide polymers, lactide/glycolide copolymers or polyoxyethylene/polyoxy-propylene copolymers. In one embodiment, the pharmaceutical composition of the present disclosure includes isotonic saline.

Pharmaceutically acceptable carriers, excipients or diluents for therapeutic use are well known in the pharmaceutical art, and are described, for example, in Remington's Pharmaceutical Sciences, Mack Publishing Co. (A. R Gennaro edit. 1985).

Pharmaceutical carriers, excipients or diluents can be selected with regard to the intended route of administration and standard pharmaceutical practice.

In one embodiment, pharmaceutical compositions described herein may be administered intravenously, intraarterially, subcutaneously, intradermally or intramuscularly. In certain embodiments, the pharmaceutical composition is formulated for local administration or systemic administration. Systemic administration may include enteral administration, which involves absorption through the gastrointestinal tract, or parenteral administration. As used herein, "parenteral administration" refers to the administration in any manner other than through the gastrointestinal tract, such as by intravenous injection. In a preferred embodiment, the pharmaceutical composition is formulated for intramuscular administration. In another embodiment, the pharmaceutical composition is formulated for systemic administration, e.g., for intravenous administration.

The term "co-administering" as used herein means a process whereby different compounds or compositions (e.g., RNA encoding different vaccine antigens) are administered to the same patient. The different compounds or compositions may be administered simultaneously, at essentially the same time, or sequentially.

Treatments

The present invention provides methods and agents for inducing an adaptive immune response against a virus, in particular different strains of a virus, in a subject comprising administering an effective amount of a composition described herein, e.g., a composition comprising a vaccine antigen or nucleic acid, e.g., RNA, encoding vaccine antigen described herein.

In one embodiment, the methods and agents described herein provide immunity in a subject to coronavirus, coronavirus infection, or to a disease or disorder associated with coronavirus. The present invention thus provides methods and agents for treating or preventing the infection, disease, or disorder associated with coronavirus.

In one embodiment, the methods and agents described herein are administered to a subject having an infection, disease, or disorder associated with a virus. In one embodiment, the methods and agents described herein are administered to a subject at risk for developing the infection, disease, or disorder associated with a virus. For example, the methods and agents described herein may be administered to a subject who is at risk for being in contact with a virus. In one embodiment, the methods and agents described herein are administered to a subject who lives in, traveled to, or is expected to travel to a geographic region in which a virus is prevalent. In one embodiment, the methods and agents described herein are administered to a subject who is in contact with or expected to be in contact with another person who lives in, traveled to, or is expected to travel to a geographic region in which a virus is prevalent. In one embodiment, the methods and agents described herein are administered to a subject who has knowingly been exposed to a virus through their occupation, or other contact. In one embodiment, a virus is a coronavirus such as SARS-CoV-2.

For a composition to be useful as a vaccine, the composition must induce an immune response against a virus antigen in a cell, tissue or subject (e.g., a human). In some embodiments, the

CA 03215771 2023-10-17

composition induces an immune response against a virus antigen in a cell, tissue or subject (e.g., a human). In some instances, the vaccine induces a protective immune response in a mammal. The therapeutic compounds or compositions of the invention may be administered prophylactically (i.e., to prevent a disease or disorder) or therapeutically (i.e., to treat a disease or disorder) to subjects suffering from, or at risk of (or susceptible to) developing a disease or disorder. Such subjects may be identified using standard clinical methods. In the context of the present invention, prophylactic administration occurs prior to the manifestation of overt clinical symptoms of disease, such that a disease or disorder is prevented or alternatively delayed in its progression. In the context of the field of medicine, the term "prevent" encompasses any activity, which reduces the burden of mortality or morbidity from disease. Prevention can occur at primary, secondary and tertiary prevention levels. While primary prevention avoids the development of a disease, secondary and tertiary levels of prevention encompass activities aimed at preventing the progression of a disease and the emergence of symptoms as well as reducing the negative impact of an already established disease by restoring function and reducing disease-related complications.

In some embodiments, administration of an immunogenic composition or vaccine of the present invention may be performed by single administration or boosted by multiple administrations.

In some embodiments, an amount the RNA described herein from 0.1 μg to 300 μg, 0.5 μg to 200 μg, or 1 μg to 100 μg, such as about 1 μg, about 3 μg, about 10 μg, about 30 μg, about 50 μg, or about 100 μg may be administered per dose.

In some embodiments, a regimen described herein includes at least one dose. In some embodiments, a regimen includes a first dose and at least one subsequent dose. In some embodiments, the first dose is the same amount as at least one subsequent dose. In some embodiments, the first dose is the same amount as all subsequent doses. In some embodiments, the first dose is a different amount as at least one subsequent dose. In some embodiments, the first dose is a different amount than all subsequent doses. In some embodiments, a regimen comprises two doses. In some embodiments, a provided regimen consists of two doses.

In one embodiment, the invention envisions administration of a single dose. In one embodiment, the invention envisions administration of a priming dose followed by one or more booster doses. The booster dose or the first booster dose may be administered 7 to 28 days or 14 to 24 days following administration of the priming dose.

In some embodiments, an amount of the RNA described herein of 60 µg or lower, 50 µg or lower, 40 µg or lower, or 30 µg or lower may be administered per dose.

In some embodiments, an amount of the RNA described herein of at least 0.25 µg, at least 0.5 µg, at least 1 µg, at least 2 µg, at least 3 µg, at least 4 µg, at least 5 µg, at least 10 µg, at least 20 µg, at least 30 µg, or at least 40 µg may be administered per dose.

In some embodiments, an amount of the RNA described herein of 0.25 µg to 60 µg, 0.5 µg to 55 µg, 1 µg to 50 µg, 5 µg to 40 µg, or 10 µg to 30 µg may be administered per dose.

If different RNA moleules are administered (e.g., encoding different polyspecific virus protein amino acid sequences), an amount or a dosis of RNA given herein may relate to the combined amount of the different RNA moleules administered. The different RNA moleules may be administered at the same time or essentially at the same time.In some embodiments, a regimen administered to a subject may be or comprise a single dose. In some embodiments, a regimen administered to a subject may comprise a plurality of doses (e.g., at least two doses, at least three doses, or more). In some embodiments, a regimen administered to a subject may comprise a first dose and a second dose, which are given at least 2 weeks apart, at least 3 weeks apart, at least 4 weeks apart, or more. In some embodiments, such doses may be at least 1 month, at least 2 months, at least 3 months, at least 4 months, at least 5 months, at least 6 months, at least 7 months, at least 8 months, at least 9 months, at least 10 months, at least 11 months, at least 12 months, or more apart. In some embodiments, doses may be administered days apart, such as 1, 2, 3, 4, 5, 6, 7, 8, 9 ,10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60 or more days apart. In some embodiments, doses may be administered about 1 to about 3 weeks apart, or about 1 to about 4 weeks apart, or about 1 to about 5 weeks apart, or about 1 to about 6 weeks apart, or about 1 to more than 6 weeks apart. In some embodiments, doses may be separated by a period of about 7 to about 60 days, such as for example about 14 to about 48 days, etc. In some

CA 03215771 2023-10-17

embodiments, a minimum number of days between doses may be about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 or more. In some embodiments, a maximum number of days between doses may be about 60, 59, 58, 57, 56, 55, 54, 53, 52, 51, 50, 49, 48, 47, 46, 45, 44, 43, 42, 41, 40, 39, 38, 37, 36, 35, 34, 33, 32, 31, 30, 29, 28, 27, 26, 25, 24, 23, 22, 21, or fewer. In some embodiments, doses may be about 21 to about 28 days apart. In some embodiments, doses may be about 19 to about 42 days apart. In some embodiments, doses may be about 7 to about 28 days apart. In some embodiments, doses may be about 14 to about 24 days. In some embodiments, doses may be about 21 to about 42 days.

In some embodiments, a vaccination regimen comprises a first dose and a second dose. In some embodiments, a first dose and a second dose are administered by at least 21 days apart. In some embodiments, a first dose and a second dose are administered by at least 28 days apart.

In some embodiments, a first dose and a second dose (and/or other subsequent dose) may be administered by intramuscular injection. In some embodiments, a first dose and a second dose may be administered in the deltoid muscle. In some embodiments, a first dose and a second dose may be administered in the same arm. In some embodiments, an mRNA composition described herein is administered (e.g., by intramuscular injection) as a series of two doses (e.g., 0.3 mL each) 21 days part. In some embodiments, each dose is about 30 ug. In some embodiments, each dose may be higher than 30 ug, e.g., about 40 ug, about 50 ug, about 60 ug. In some embodiments, each dose may be lower than 30 ug, e.g., about 20 ug, about 10 ug, about 5 ug, etc. In some embodiments, each dose is about 3 ug or lower, e.g., about 1 ug. In some such embodiments, an mRNA composition described herein is administered to subjects of age 16 or older (including, e.g., 16-85 years). In some such embodiments, an mRNA composition described herein is administered to subjects of age 18-55. In some such embodiments, an mRNA composition escribed herein is administered to subjects of age 56-85. In some embodiments, an mRNA composition described herein is administered (e.g., by intramuscular injection) as a single dose.

In one embodiment, an amount of the RNA described herein of about 30 μg is administered per dose. In one embodiment, at least two of such doses are administered. For example, a second dose may be administered about 21 days following administration of the first dose.

In some embodiments, the efficacy of the RNA vaccine described herein (e.g., administered in two doses, wherein a second dose may be administered about 21 days following administration of the first dose, and administered, for example, in an amount of about 30 μg per dose) is at least 70%, at least 80%, at least 90, or at least 95% beginning 7 days after administration of the second dose (e.g., beginning 28 days after administration of the first dose if a second dose is administered 21 days following administration of the first dose). In some embodiments, such efficacy is observed in populations of age of at least 50, at least 55, at least 60, at least 65, at least 70, or older. In some embodiments, the efficacy of the RNA vaccine described herein (e.g., administered in two doses, wherein a second dose may be administered about 21 days following administration of the first dose, and administered, for example, in an amount of about 30 μg per dose) beginning 7 days after administration of the second dose (e.g., beginning 28 days after administration of the first dose if a second dose is administered 21 days following administration of the first dose) in populations of age of at least 65, such as 65 to 80, 65 to 75, or 65 to 70, is at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, or at least 95%. Such efficacy may be observed over time periods of up to 1 month, 2 months, 3 months, 6 months or even longer.

In some embodiments, compositions and/or methods described herein are characterized in that 7 days after administration of the second dose, the protective efficacy is at least 60%, e.g., at least 70%, at least 80%, at least 90, or at least 95%. In one embodiment, compositions and/or methods described herein are characterized in that 7 days after administration of the second dose, the protective efficacy is at least 70%. In one embodiment, compositions and/or methods described herein are characterized in that 7 days after administration of the second dose, the protective efficacy is at least 80%. In one embodiment, compositions and/or methods described herein are characterized in that 7 days after administration of the second dose, the protective efficacy is at least 90%. In one embodiment, compositions and/or methods described herein are characterized in that 7 days after administration of the second dose, the protective efficacy is at least 95%.

In one embodiment, vaccine efficacy is defined as the percent reduction in the number of subjects with evidence of infection (vaccinated subjects vs. non-vaccinated subjects).

CA 03215771 2023-10-17

In one embodiment, efficacy is assessed through surveillance for potential cases of COVID-19. If, at any time, a patient develops acute respiratory illness, for the purposes herein, the patient can be considered to potentially have COVID-19 illness. The assessments can include a nasal (midturbinate) swab, which may be tested using a reverse transcription-polymerase chain reaction (RT-PCR) test to detect SARS-CoV-2. In addition, clinical information and results from local standard-of-care tests can be assessed.

In some embodiments, efficacy assessments may utilize a definition of SARS-CoV-2-related cases wherein:

• Confirmed COVID-19: presence of at least 1 of the following symptoms and SARS-CoV-2 NAAT (nucleic acid amplification-based test) positive during, or within 4 days before or after, the symptomatic period: fever; new or increased cough; new or increased shortness of breath; chills; new or increased muscle pain; new loss of taste or smell; sore throat; diarrhea; vomiting.

Alternatively or additionally, in some embodiments, efficacy assessments may utilize a definition of SARS-CoV-2-related cases wherein one or more of the following additional symptoms defined by the CDC can be considered: fatigue; headache; nasal congestion or runny nose; nausea.

In some embodiments, efficacy assessments may utilize a definition of SARS-CoV-2-related severe cases

• Confirmed severe COVID-19: confirmed COVID-19 and presence of at least 1 of the following: clinical signs at rest indicative of severe systemic illness (e.g., RR ≥30 breaths per minute, HR ≥125 beats per minute, $SpO_2 \leq 93\%$ on room air at sea level, or $PaO_2/FiO_2 < 300$mm Hg); respiratory failure (which can be defined as needing high-flow oxygen, noninvasive ventilation, mechanical ventilation, or ECMO); evidence of shock (e.g., SBP <90 mm Hg, DBP <60 mm Hg, or requiring vasopressors); significant acute renal, hepatic, or neurologic dysfunction; admission to an ICU; death.

Alternatively or additionally, in some embodiments a serological definition can be used for patients without clinical presentation of COVID-19: e.g., confirmed seroconversion to SARS-CoV-2 without confirmed COVID-19: e.g., positive N-binding antibody result in a patient with a prior negative N-binding antibody result.

CA 03215771 2023-10-17

In some embodiments, any or all of the following assays can be performed on serum samples: SARS-CoV-2 neutralization assay; S1-binding IgG level assay; RBD-binding IgG level assay; N-binding antibody assay.

In one embodiment, methods and agents described herein are administered to a paediatric population. In various embodiments, the paediatric population comprises or consists of subjects under 18 years, e.g., 5 to less than 18 years of age, 12 to less than 18 years of age, 16 to less than 18 years of age, 12 to less than 16 years of age, or 5 to less than 12 years of age. In various embodiments, the paediatric population comprises or consists of subjects under 5 years, e.g., 2 to less than 5 years of age, 12 to less than 24 months of age, 7 to less than 12 months of age, or less than 6 months of age.

In one embodiment, the paediatric population comprises or consists of subjects 12 to less than 18 years of age including subjects 16 to less than 18 years of age and/or subjects 12 to less than 16 years of age. In this embodiment, treatments may comprise 2 vaccinations 21 days apart, wherein, in one embodiment, the vaccine is administered in an amount of 30 µg RNA per dose, e.g., by intramuscular administration.

In one embodiment, the paediatric population comprises or consists of subjects 5 to less than 18 years of age including subjects 12 to less than 18 years of age and/or subjects 5 to less than 12 years of age. In this embodiment, treatments may comprise 2 vaccinations 21 days apart, wherein, in various embodiments, the vaccine is administered in an amount of 10 µg, 20µg, or 30 µg RNA per dose, e.g., by intramuscular administration.

In one embodiment, the paediatric population comprises or consists of subjects less than 5 years of age including subjects 2 to less than 5 years of age, subjects 12 to less than 24 months of age, subjects 7 to less than 12 months of age, subjects 6 to less than 12 months of age and/or subjects less than 6 months of age. In this embodiment, treatments may comprise 2 vaccinations, e.g., 21 to 42 days apart, e.g., 21 days apart, wherein, in various embodiments, the vaccine is administered in an amount of 10 µg, 20µg, or 30 µg RNA per dose, e.g., by intramuscular administration.

In some embodiments, populations to be treated with RNA described herein comprise, essentially consist of, or consist of subjects of age of at least 50, at least 55, at least 60, or at least 65. In some embodiments, populations to be treated with RNA described herein

CA 03215771 2023-10-17

comprise, essentially consist of, or consist of subjects of age of between 55 to 90, 60 to 85, or 65 to 85.

In some embodiments, the period of time between the doses administered is at least 7 days, at least 14 days, or at least 21 days. In some embodiments, the period of time between the doses administered is between 7 days and 28 days such as between 14 days and 23 days.

In some embodiments, no more than 5 doses, no more than 4 doses, or no more than 3 doses of the RNA described herein may be administered to a subject.

In some embodiments, the methods and agents described herein provide a neutralizing effect in a subject to coronavirus, coronavirus infection, or to a disease or disorder associated with coronavirus.

In some embodiments, the methods and agents described herein following administration to a subject induce an immune response that blocks or neutralizes coronavirus in the subject. In some embodiments, the methods and agents described herein following administration to a subject induce the generation of antibodies such as IgG antibodies that block or neutralize coronavirus in the subject. In some embodiments, the methods and agents described herein following administration to a subject induce an immune response that blocks or neutralizes coronavirus S protein binding to ACE2 in the subject. In some embodiments, the methods and agents described herein following administration to a subject induce the generation of antibodies that block or neutralize coronavirus S protein binding to ACE2 in the subject.

As used herein, the term "neutralization" refers to an event in which binding agents such as antibodies bind to a biological active site of a virus such as a receptor binding protein, thereby inhibiting the viral infection of cells. As used herein, the term "neutralization" with respect to coronavirus, in particular coronavirus S protein, refers to an event in which binding agents such as antibodies bind to the RBD domain of the S protein, thereby inhibiting the viral infection of cells. In particular, the term "neutralization" refers to an event in which binding agents eliminate or significantly reduce virulence (e.g. ability of infecting cells) of viruses of interest.

The type of immune response generated in response to an antigenic challenge can generally be distinguished by the subset of T helper (Th) cells involved in the response. Immune responses can be broadly divided into two types: Th1 and Th2. Th1 immune activation is

CA 03215771 2023-10-17

optimized for intracellular infections such as viruses, whereas Th2 immune responses are optimized for humoral (antibody) responses. Th1 cells produce interleukin 2 (IL-2), tumor necrosis factor (TNFα) and interferon gamma (IFNγ). Th2 cells produce IL-4, IL-5, IL-6, IL-9, IL-10 and IL-13. Th1 immune activation is the most highly desired in many clinical situations. Vaccine compositions specialized in eliciting Th2 or humoral immune responses are generally not effective against most viral diseases.

In some embodiments, the methods and agents described herein following administration to a subject induce or promote a Th1-mediated immune response in the subject. In some embodiments, the methods and agents described herein following administration to a subject induce or promote a cytokine profile that is typical for a Th1-mediated immune response in the subject. In some embodiments, the methods and agents described herein following administration to a subject induce or promote the production of interleukin 2 (IL-2), tumor necrosis factor (TNFα) and/or interferon gamma (IFNγ) in the subject. In some embodiments, the methods and agents described herein following administration to a subject induce or promote the production of interleukin 2 (IL-2) and interferon gamma (IFNγ) in the subject. In some embodiments, the methods and agents described herein following administration to a subject do not induce or promote a Th2-mediated immune response in the subject, or induce or promote a Th2-mediated immune response in the subject to a significant lower extent compared to the induction or promotion of a Th1-mediated immune response. In some embodiments, the methods and agents described herein following administration to a subject do not induce or promote a cytokine profile that is typical for a Th2-mediated immune response in the subject, or induce or promote a cytokine profile that is typical for a Th2-mediated immune response in the subject to a significant lower extent compared to the induction or promotion of a cytokine profile that is typical for a Th1-mediated immune response. In some embodiments, the methods and agents described herein following administration to a subject do not induce or promote the production of IL-4, IL-5, IL-6, IL-9, IL-10 and/or IL-13, or induce or promote the production of IL-4, IL-5, IL-6, IL-9, IL-10 and/or IL-13 in the subject to a significant lower extent compared to the induction or promotion of interleukin 2 (IL-2), tumor necrosis factor (TNFα) and/or interferon gamma (IFNγ) in the subject. In some embodiments, the methods and agents described herein following

CA 03215771 2023-10-17

administration to a subject do not induce or promote the production of IL-4, or induce or promote the production of IL-4 in the subject to a significant lower extent compared to the induction or promotion of interleukin 2 (IL-2) and interferon gamma (IFNγ) in the subject.

In some embodiments, the methods and agents described herein following administration to a subject induce an antibody response, in particular a neutralizing antibody response, in the subject that targets a panel of different S protein variants such as SARS-CoV-2 S protein variants, in particular naturally occurring S protein variants. In some embodiments, the panel of different S protein variants comprises at least 5, at least 10, at least 15, or even more S protein variants. In some embodiments, such S protein variants comprise variants having amino acid modifications in the RBD domain and/or variants having amino acid modifications outside the RBD domain, e.g., in the NTD domain. Such amino acid modifications are described herein.

In one embodiment, vaccination described herein, e.g., using RNA described herein which may be administered in the amounts and regimens described herein, e.g., at two doses of 30 µg per dose e.g. administered 21 days apart, may be repated after a certain period of time, e.g., once it is observed that protection against virus infection diminishes, using the same or a different vaccine as used for the first vaccination. Such certain period of time may be at least 6 months, 1 year, two years etc. In one embodiment, the same RNA as used for the first vaccination is used for the second or further vaccination, however, at a lower dose or a lower frequency of administration. For example, the first vaccination may comprise vaccination using a dose of about 30 µg per dose, wherein in one embodiment, at least two of such doses are administered, (for example, a second dose may be administered about 21 days following administration of the first dose) and the second or further vaccination may comprise vaccination using a dose of less than about 30 µg per dose, wherein in one embodiment, only one of such doses is administered. In one embodiment, a different RNA as used for the first vaccination is used for the second or further vaccination. In one embodiment, BNT162b2 is used for the first vaccination and a vaccine antigen as describd herein with modifications present in strains prevalent at the time of the second or further vaccination is used for the second or further vaccination. In one embodiment, a vaccine antigen with modifications present in strains prevalent at the time of the first vaccination is used for the first vaccination

CA 03215771 2023-10-17

and a vaccine antigen with modifications present in strains prevalent at the time of the second or further vaccination is used for the second or further vaccination.

In one embodiment, the vaccination regimen comprises a first vaccination using at least two doses of the RNA described herein, e.g., two doses of the RNA described herein (wherein the second dose may be administered about 21 days following administration of the first dose), and a second vaccination using a single dose or multiple doses, e.g., two doses, of the RNA described herein. In various embodiments, the second vaccination is administered 3 to 24 months, 6 to 18 months, 6 to 12 months, or 5 to 7 months after administration of the first vaccination, e.g., after the initial two-dose regimen. The amount of RNA used in each dose of the second vaccination may be equal or different to the amount of RNA used in each dose of the first vaccination. In one embodiment, the amount of RNA used in each dose of the second vaccination is equal to the amount of RNA used in each dose of the first vaccination. In one embodiment, the amount of RNA used in each dose of the second vaccination and the amount of RNA used in each dose of the first vaccination is about 30 μg per dose. In one embodiment, the same RNA as used for the first vaccination is used for the second vaccination.

In one embodiment, the second vaccination results in a boosting of the immune response.

In one embodiment, the RNA described herein is co-administered with other vaccines. In some embodiments, an RNA composition described herein is co-administered with an influenza vaccine. In some embodiments, an RNA composition provided herein and other injectable vaccine(s) are administered at different times. In some embodiments, an RNA composition provided herein is administered at the same time as other injectable vaccine(s). In some such embodiments, an RNA composition provided herein and at least one another injectable vaccine(s) are administered at different injection sites.

The term "disease" refers to an abnormal condition that affects the body of an individual. A disease is often construed as a medical condition associated with specific symptoms and signs. A disease may be caused by factors originally from an external source, such as infectious disease, or it may be caused by internal dysfunctions, such as autoimmune diseases. In humans, "disease" is often used more broadly to refer to any condition that causes pain, dysfunction, distress, social problems, or death to the individual afflicted, or similar problems for those in contact with the individual. In this broader sense, it sometimes includes injuries,

CA 03215771 2023-10-17

disabilities, disorders, syndromes, infections, isolated symptoms, deviant behaviors, and atypical variations of structure and function, while in other contexts and for other purposes these may be considered distinguishable categories. Diseases usually affect individuals not only physically, but also emotionally, as contracting and living with many diseases can alter one's perspective on life, and one's personality.

In the present context, the term "treatment", "treating" or "therapeutic intervention" relates to the management and care of a subject for the purpose of combating a condition such as a disease or disorder. The term is intended to include the full spectrum of treatments for a given condition from which the subject is suffering, such as administration of the therapeutically effective compound to alleviate the symptoms or complications, to delay the progression of the disease, disorder or condition, to alleviate or relief the symptoms and complications, and/or to cure or eliminate the disease, disorder or condition as well as to prevent the condition, wherein prevention is to be understood as the management and care of an individual for the purpose of combating the disease, condition or disorder and includes the administration of the active compounds to prevent the onset of the symptoms or complications.

The term "therapeutic treatment" relates to any treatment which improves the health status and/or prolongs (increases) the lifespan of an individual. Said treatment may eliminate the disease in an individual, arrest or slow the development of a disease in an individual, inhibit or slow the development of a disease in an individual, decrease the frequency or severity of symptoms in an individual, and/or decrease the recurrence in an individual who currently has or who previously has had a disease.

The terms "prophylactic treatment" or "preventive treatment" relate to any treatment that is intended to prevent a disease from occurring in an individual. The terms "prophylactic treatment" or "preventive treatment" are used herein interchangeably.

The terms "individual" and "subject" are used herein interchangeably. They refer to a human or another mammal (e.g. mouse, rat, rabbit, dog, cat, cattle, swine, sheep, horse or primate) that can be afflicted with or is susceptible to a disease or disorder but may or may not have the disease or disorder. In many embodiments, the individual is a human being. Unless otherwise stated, the terms "individual" and "subject" do not denote a particular age, and

CA 03215771 2023-10-17

thus encompass adults, elderlies, children, and newborns. In some embodiments, subjects described herein are younger subjects (e.g., less than 25 years old, 20 years old, 18 years old, 15 years, 10 years old, or lower); alternatively or additionally, in some embodiments, subjects described herein are elderly subjects (e.g., over 55 years old, 60 years old, 65 years old, 70 years old, 75 years old, 80 years old, 85 years old, or higher).

In embodiments of the present disclosure, the "individual" or "subject" is a "patient".

The term "patient" means an individual or subject for treatment, in particular a diseased individual or subject.

In one embodiment of the disclosure, the aim is to provide an immune response against a virus from which at least a fragment of a virus protein is derived, such as coronavirus, and to prevent or treat virus infection, such as coronavirus infection.

A pharmaceutical composition comprising RNA encoding a peptide or protein comprising an epitope may be administered to a subject to elicit an immune response against an antigen comprising said epitope in the subject which may be therapeutic or partially or fully protective.

A person skilled in the art will know that one of the principles of immunotherapy and vaccination is based on the fact that an immunoprotective reaction to a disease is produced by immunizing a subject with an antigen or an epitope, which is immunologically relevant with respect to the disease to be treated. Accordingly, pharmaceutical compositions described herein are applicable for inducing or enhancing an immune response. Pharmaceutical compositions described herein are thus useful in a prophylactic and/or therapeutic treatment of a disease involving an antigen or epitope.

RNA constructs described herein encoding polypeptides comprising a modified amino acid sequence of at least a portion of a SARS-CoV-2 S protein, for example at least an RBD portion of a SARS-CoV-2 S protein, or a full-length or essentially full-length SARS-CoV-2-encoded S protein may be particularly useful and/or effective for use as or in an immunogenic composition (e.g., a vaccine), and/or for achieving immunological effects as described herein (e.g., generation of SARS-CoV-2 neutralizing antibodies, and/or T cell responses (e.g., CD4+ and/or CD8+ T cell responses)).

As used herein, "immune response" refers to an integrated bodily response to an antigen or a cell expressing an antigen and refers to a cellular immune response and/or a humoral immune

CA 03215771 2023-10-17

response. The immune system is divided into a more primitive innate immune system, and acquired or adaptive immune system of vertebrates, each of which contains humoral and cellular components.

"Cell-mediated immunity", "cellular immunity", "cellular immune response", or similar terms are meant to include a cellular response directed to cells characterized by expression of an antigen, in particular characterized by presentation of an antigen with class I or class II MHC. The cellular response relates to immune effector cells, in particular to cells called T cells or T lymphocytes which act as either "helpers" or "killers". The helper T cells (also termed $CD4^+$ T cells) play a central role by regulating the immune response and the killer cells (also termed cytotoxic T cells, cytolytic T cells, $CD8^+$ T cells or CTLs) kill diseased cells such as virus-infected cells, preventing the production of more diseased cells.

An immune effector cell includes any cell which is responsive to vaccine antigen. Such responsiveness includes activation, differentiation, proliferation, survival and/or indication of one or more immune effector functions. The cells include, in particular, cells with lytic potential, in particular lymphoid cells, and are preferably T cells, in particular cytotoxic lymphocytes, preferably selected from cytotoxic T cells, natural killer (NK) cells, and lymphokine-activated killer (LAK) cells. Upon activation, each of these cytotoxic lymphocytes triggers the destruction of target cells. For example, cytotoxic T cells trigger the destruction of target cells by either or both of the following means. First, upon activation T cells release cytotoxins such as perforin, granzymes, and granulysin. Perforin and granulysin create pores in the target cell, and granzymes enter the cell and trigger a caspase cascade in the cytoplasm that induces apoptosis (programmed cell death) of the cell. Second, apoptosis can be induced via Fas-Fas ligand interaction between the T cells and target cells.

The term "effector functions" in the context of the present invention includes any functions mediated by components of the immune system that result, for example, in the neutralization of a pathogenic agent such as a virus and/or in the killing of diseased cells such as virus-infected cells. In one embodiment, the effector functions in the context of the present invention are T cell mediated effector functions. Such functions comprise in the case of a helper T cell ($CD4^+$ T cell) the release of cytokines and/or the activation of $CD8^+$ lymphocytes (CTLs) and/or B cells, and in the case of CTL the elimination of cells, i.e., cells characterized by

CA 03215771 2023-10-17

expression of an antigen, for example, via apoptosis or perforin-mediated cell lysis, production of cytokines such as IFN-γ and TNF-α, and specific cytolytic killing of antigen expressing target cells.

The term "immune effector cell" or "immunoreactive cell" in the context of the present invention relates to a cell which exerts effector functions during an immune reaction. An "immune effector cell" in one embodiment is capable of binding an antigen such as an antigen presented in the context of MHC on a cell or expressed on the surface of a cell and mediating an immune response. For example, immune effector cells comprise T cells (cytotoxic T cells, helper T cells, tumor infiltrating T cells), B cells, natural killer cells, neutrophils, macrophages, and dendritic cells. Preferably, in the context of the present invention, "immune effector cells" are T cells, preferably $CD4^+$ and/or $CD8^+$ T cells, most preferably $CD8^+$ T cells. According to the invention, the term "immune effector cell" also includes a cell which can mature into an immune cell (such as T cell, in particular T helper cell, or cytolytic T cell) with suitable stimulation. Immune effector cells comprise $CD34^+$ hematopoietic stem cells, immature and mature T cells and immature and mature B cells. The differentiation of T cell precursors into a cytolytic T cell, when exposed to an antigen, is similar to clonal selection of the immune system.

A "lymphoid cell" is a cell which is capable of producing an immune response such as a cellular immune response, or a precursor cell of such cell, and includes lymphocytes, preferably T lymphocytes, lymphoblasts, and plasma cells. A lymphoid cell may be an immune effector cell as described herein. A preferred lymphoid cell is a T cell.

The terms "T cell" and "T lymphocyte" are used interchangeably herein and include T helper cells (CD4+ T cells) and cytotoxic T cells (CTLs, CD8+ T cells) which comprise cytolytic T cells. The term "antigen-specific T cell" or similar terms relate to a T cell which recognizes the antigen to which the T cell is targeted and preferably exerts effector functions of T cells.

T cells belong to a group of white blood cells known as lymphocytes, and play a central role in cell-mediated immunity. They can be distinguished from other lymphocyte types, such as B cells and natural killer cells by the presence of a special receptor on their cell surface called T cell receptor (TCR). The thymus is the principal organ responsible for the maturation of T cells. Several different subsets of T cells have been discovered, each with a distinct function.

CA 03215771 2023-10-17

T helper cells assist other white blood cells in immunologic processes, including maturation of B cells into plasma cells and activation of cytotoxic T cells and macrophages, among other functions. These cells are also known as CD4+ T cells because they express the CD4 glycoprotein on their surface. Helper T cells become activated when they are presented with peptide antigens by MHC class II molecules that are expressed on the surface of antigen presenting cells (APCs). Once activated, they divide rapidly and secrete small proteins called cytokines that regulate or assist in the active immune response.

Cytotoxic T cells destroy virally infected cells and tumor cells, and are also implicated in transplant rejection. These cells are also known as CD8+ T cells since they express the CD8 glycoprotein on their surface. These cells recognize their targets by binding to antigen associated with MHC class I, which is present on the surface of nearly every cell of the body.

A majority of T cells have a T cell receptor (TCR) existing as a complex of several proteins. The TCR of a T cell is able to interact with immunogenic peptides (epitopes) bound to major histocompatibility complex (MHC) molecules and presented on the surface of target cells. Specific binding of the TCR triggers a signal cascade inside the T cell leading to proliferation and differentiation into a maturated effector T cell. The actual T cell receptor is composed of two separate peptide chains, which are produced from the independent T cell receptor alpha and beta (TCRα and TCRβ) genes and are called α- and β-TCR chains. γδ T cells (gamma delta T cells) represent a small subset of T cells that possess a distinct T cell receptor (TCR) on their surface. However, in γδ T cells, the TCR is made up of one γ-chain and one δ-chain. This group of T cells is much less common (2% of total T cells) than the αβ T cells.

"Humoral immunity" or "humoral immune response" is the aspect of immunity that is mediated by macromolecules found in extracellular fluids such as secreted antibodies, complement proteins, and certain antimicrobial peptides. It contrasts with cell-mediated immunity. Its aspects involving antibodies are often called antibody-mediated immunity.

Humoral immunity refers to antibody production and the accessory processes that accompany it, including: Th2 activation and cytokine production, germinal center formation and isotype switching, affinity maturation and memory cell generation. It also refers to the effector functions of antibodies, which include pathogen neutralization, classical complement activation, and opsonin promotion of phagocytosis and pathogen elimination.

CA 03215771 2023-10-17

In humoral immune response, first the B cells mature in the bone marrow and gain B-cell receptors (BCR's) which are displayed in large number on the cell surface. These membrane-bound protein complexes have antibodies which are specific for antigen detection. Each B cell has a unique antibody that binds with an antigen. The mature B cells migrate from the bone marrow to the lymph nodes or other lymphatic organs, where they begin to encounter pathogens. When a B cell encounters an antigen, the antigen is bound to the receptor and taken inside the B cell by endocytosis. The antigen is processed and presented on the B cell's surface again by MHC-II proteins. The B cell waits for a helper T cell (TH) to bind to the complex. This binding will activate the TH cell, which then releases cytokines that induce B cells to divide rapidly, making thousands of identical clones of the B cell. These daughter cells either become plasma cells or memory cells. The memory B cells remain inactive here; later when these memory B cells encounter the same antigen due to reinfection, they divide and form plasma cells. On the other hand, the plasma cells produce a large number of antibodies which are released free into the circulatory system. These antibodies will encounter antigens and bind with them. This will either interfere with the chemical interaction between host and foreign cells, or they may form bridges between their antigenic sites hindering their proper functioning, or their presence will attract macrophages or killer cells to attack and phagocytose them.

The term "antibody" includes an immunoglobulin comprising at least two heavy (H) chains and two light (L) chains inter-connected by disulfide bonds. Each heavy chain is comprised of a heavy chain variable region (abbreviated herein as VH) and a heavy chain constant region. Each light chain is comprised of a light chain variable region (abbreviated herein as VL) and a light chain constant region. The VH and VL regions can be further subdivided into regions of hypervariability, termed complementarity determining regions (CDR), interspersed with regions that are more conserved, termed framework regions (FR). Each VH and VL is composed of three CDRs and four FRs, arranged from amino-terminus to carboxy-terminus in the following order: FR1, CDR1, FR2, CDR2, FR3, CDR3, FR4. The variable regions of the heavy and light chains contain a binding domain that interacts with an antigen. The constant regions of the antibodies may mediate the binding of the immunoglobulin to host tissues or factors, including various cells of the immune system (e.g., effector cells) and the first component (Clq)

CA 03215771 2023-10-17

of the classical complement system. An antibody binds, preferably specifically binds with an antigen.

Antibodies expressed by B cells are sometimes referred to as the BCR (B cell receptor) or antigen receptor. The five members included in this class of proteins are IgA, IgG, IgM, IgD, and IgE. IgA is the primary antibody that is present in body secretions, such as saliva, tears, breast milk, gastrointestinal secretions and mucus secretions of the respiratory and genitourinary tracts. IgG is the most common circulating antibody. IgM is the main immunoglobulin produced in the primary immune response in most subjects. It is the most efficient immunoglobulin in agglutination, complement fixation, and other antibody responses, and is important in defense against bacteria and viruses. IgD is the immunoglobulin that has no known antibody function, but may serve as an antigen receptor. IgE is the immunoglobulin that mediates immediate hypersensitivity by causing release of mediators from mast cells and basophils upon exposure to allergen.

An "antibody heavy chain", as used herein, refers to the larger of the two types of polypeptide chains present in antibody molecules in their naturally occurring conformations.

An "antibody light chain", as used herein, refers to the smaller of the two types of polypeptide chains present in antibody molecules in their naturally occurring conformations, κ and λ light chains refer to the two major antibody light chain isotypes.

The present disclosure contemplates an immune response that may be protective, preventive, prophylactic and/or therapeutic. As used herein, "induces [or inducing] an immune response" may indicate that no immune response against a particular antigen was present before induction or it may indicate that there was a basal level of immune response against a particular antigen before induction, which was enhanced after induction. Therefore, "induces [or inducing] an immune response" includes "enhances [or enhancing] an immune response".

The term "immunotherapy" relates to the treatment of a disease or condition by inducing, or enhancing an immune response. The term "immunotherapy" includes antigen immunization or antigen vaccination.

The terms "immunization" or "vaccination" describe the process of administering an antigen to an individual with the purpose of inducing an immune response, for example, for therapeutic or prophylactic reasons.

CA 03215771 2023-10-17

The term "macrophage" refers to a subgroup of phagocytic cells produced by the differentiation of monocytes. Macrophages which are activated by inflammation, immune cytokines or microbial products nonspecifically engulf and kill foreign pathogens within the macrophage by hydrolytic and oxidative attack resulting in degradation of the pathogen. Peptides from degraded proteins are displayed on the macrophage cell surface where they can be recognized by T cells, and they can directly interact with antibodies on the B cell surface, resulting in T and B cell activation and further stimulation of the immune response. Macrophages belong to the class of antigen presenting cells. In one embodiment, the macrophages are splenic macrophages.

The term "dendritic cell" (DC) refers to another subtype of phagocytic cells belonging to the class of antigen presenting cells. In one embodiment, dendritic cells are derived from hematopoietic bone marrow progenitor cells. These progenitor cells initially transform into immature dendritic cells. These immature cells are characterized by high phagocytic activity and low T cell activation potential. Immature dendritic cells constantly sample the surrounding environment for pathogens such as viruses and bacteria. Once they have come into contact with a presentable antigen, they become activated into mature dendritic cells and begin to migrate to the spleen or to the lymph node. Immature dendritic cells phagocytose pathogens and degrade their proteins into small pieces and upon maturation present those fragments at their cell surface using MHC molecules. Simultaneously, they upregulate cell-surface receptors that act as co-receptors in T cell activation such as CD80, CD86, and CD40 greatly enhancing their ability to activate T cells. They also upregulate CCR7, a chemotactic receptor that induces the dendritic cell to travel through the blood stream to the spleen or through the lymphatic system to a lymph node. Here they act as antigen-presenting cells and activate helper T cells and killer T cells as well as B cells by presenting them antigens, alongside non-antigen specific co-stimulatory signals. Thus, dendritic cells can actively induce a T cell- or B cell-related immune response. In one embodiment, the dendritic cells are splenic dendritic cells.

The term "antigen presenting cell" (APC) is a cell of a variety of cells capable of displaying, acquiring, and/or presenting at least one antigen or antigenic fragment on (or at) its cell surface. Antigen-presenting cells can be distinguished in professional antigen presenting cells and non-professional antigen presenting cells.

CA 03215771 2023-10-17

The term "professional antigen presenting cells" relates to antigen presenting cells which constitutively express the Major Histocompatibility Complex class II (MHC class II) molecules required for interaction with naive T cells. If a T cell interacts with the MHC class II molecule complex on the membrane of the antigen presenting cell, the antigen presenting cell produces a co-stimulatory molecule inducing activation of the T cell. Professional antigen presenting cells comprise dendritic cells and macrophages.

The term "non-professional antigen presenting cells" relates to antigen presenting cells which do not constitutively express MHC class II molecules, but upon stimulation by certain cytokines such as interferon-gamma. Exemplary, non-professional antigen presenting cells include fibroblasts, thymic epithelial cells, thyroid epithelial cells, glial cells, pancreatic beta cells or vascular endothelial cells.

"Antigen processing" refers to the degradation of an antigen into procession products, which are fragments of said antigen (e.g., the degradation of a protein into peptides) and the association of one or more of these fragments (e.g., via binding) with MHC molecules for presentation by cells, such as antigen presenting cells to specific T cells.

The term "disease involving an antigen" refers to any disease which implicates an antigen, e.g. a disease which is characterized by the presence of an antigen. The disease involving an antigen can be an infectious disease. As mentioned above, the antigen may be a disease-associated antigen, such as a viral antigen. In one embodiment, a disease involving an antigen is a disease involving cells expressing an antigen, preferably on the cell surface.

The term "infectious disease" refers to any disease which can be transmitted from individual to individual or from organism to organism, and is caused by a microbial agent (e.g. common cold). Infectious diseases are known in the art and include, for example, a viral disease, a bacterial disease, or a parasitic disease, which diseases are caused by a virus, a bacterium, and a parasite, respectively. In this regard, the infectious disease can be, for example, hepatitis, sexually transmitted diseases (e.g. chlamydia or gonorrhea), tuberculosis, HIV/acquired immune deficiency syndrome (AIDS), diphtheria, hepatitis B, hepatitis C, cholera, severe acute respiratory syndrome (SARS), the bird flu, and influenza.

In one embodiment of the present disclosure, a virus is an RNA virus, in particular an RNA virus causing an infectious disease.

CA 03215771 2023-10-17

Citation of documents and studies referenced herein is not intended as an admission that any of the foregoing is pertinent prior art. All statements as to the contents of these documents are based on the information available to the applicants and do not constitute any admission as to the correctness of the contents of these documents.

The following description is presented to enable a person of ordinary skill in the art to make and use the various embodiments. Descriptions of specific devices, techniques, and applications are provided only as examples. Various modifications to the examples described herein will be readily apparent to those of ordinary skill in the art, and the general principles defined herein may be applied to other examples and applications without departing from the spirit and scope of the various embodiments. Thus, the various embodiments are not intended to be limited to the examples described herein and shown, but are to be accorded the scope consistent with the claims.

Examples

Example 1: Bivalent vaccine design

In order to generate broadly neutralizing humoral immunity against multiple established and evolving SARS-CoV-2 variant strains of concern that may increase transmissibility of the virus or decrease efficacy of current vaccines effective against the prototypic Wuhan strain, a novel bivalent vaccine approach was established. The basis of the bivalent vaccine design forms the spike (S) glycoprotein sequence of the fast-spreading lineage B.1.1.7 [a.k.a. Variant of Concern 202012/01 (VOC-202012/01) or 20B/501Y.V1], as the B.1.1.7 variant is on its way to become the dominant SARS-CoV-2 variant worldwide. The B.1.1.7 variant contains several mutations in the S glycoprotein (Table 4) and has been shown to be inherently more transmissible, with a growth rate that has been estimated to be 40-70% higher than other SARS-CoV-2 lineages (Volz et al., Nature, 2021; Washington et al., Cell, 2021). Spike mutations to the SARS-CoV-2 S receptor-binding domain (RBD) and/or the N-terminal domain (NTD) found in circulating variants of concern or variants of interest, *i.e.* lineages B.1.351 (a.k.a. 20H/501Y.V2), P.1, B.1.427/B.1.429 (a.k.a. CAL.20C), B.1.526 (Table 4), as well as other mutations with high prevalence that have been shown to confer immune escape to either neutralizing monoclonal antibodies (nAb) or human COVID-19 convalescent serum were introduced and distributed among two B.1.1.7 S sequences (Table 5, Figure 6). On the one hand, the important mutational cluster K417N, E484K and N501Y, which is found in the B.1.351 lineage, was kept conserved. On the other hand, single point mutations were introduced in a way that amino acid exchanges are a reasonable distance apart and presumably result in discrete conformational epitopes. For example, the S477N mutation that is located in the binding epitope of class 1 RBD-targeted nAbs was combined with the L452R and N439K mutations, that are located in either the binding epitope of class 2 or class 3 RBD-targeted nAbs (Barnes et al., Nature, 2020). For the NTD, only additional mutations targeting surface exposed amino acid residues were considered, such as L18F, D80A, D215G, R246I and D253G. Again, these mutations were introduced in a way that amino acid exchanges are a reasonable distance apart and presumably result in discrete conformational epitopes.

CA 03215771 2023-10-17

Table 4: Mutations found in the spike glycoprotein of SARS-CoV-2 variants of concern and variants with high prevalence. For lineage B.1.526 two sub-lineages with different key amino acid exchanges in the RBD are known. del = deletion

Lineage	Amino acid substitution
B.1.1.7	del69-70, del144, N501Y, A570D, D614G, P681H, T716I, S982A, D1118H
B.1.351	L18F, D80A, D215G, del242-244, R246I, K417N, E484K, N501Y, D614G, A701V
P.1	L18F, T20N, P26S, D138Y, R190S, K417T, E484K, N501Y, D614G, H655Y, T1027I, V1176F
B.1.427/ B.1.429	S13I, W152C, L452R, D614G
B.1.526	L5F, T95I, D253G, S477N or E484K, D614G, A701V
B.1.160	S477N, D614G
B.1.258	del69-70, N439K, D614G, N679S, A845S

Table 5: Exemplary novel bivalent vaccine design based on the B.1.1.7 lineage spike protein. Additional non-B.1.1.7 lineage spike mutations are indicated in bold.

Mutated Spike sequence	Amino acid substitution
S-Seq1	**L18F**, del69-70, del144, **D253G**, **N439K**, **L452R**, **S477N**, N501Y, A570D, D614G, P681H, **A701V**, T716I, S982A, D1118H
S-Seq2	del69-70, **D80A**, del144, **D215G**, **R246I**, **K417N**, **E484K**, N501Y, A570D, D614G, P681H, T716I, S982A, D1118H

Example 2: In vitro testing

HEK293T cells were transfected in triplicates with 0.15 µg/mL modRNA using a commercial transfection reagent or 0.15 µg/mL LNP-formulated modRNA encoding vaccine candidates BNT162b2, BNT162b2(Alpha), BNT162b2(Alpha+SA) and BNT162b2(Alpha;L452R+E484Q). For sequence specification of the constructs see **Table 6**. After transfection, cells were incubated to express the vaccine candidate encoded S protein for 18h. Afterwards, cells were harvested

CA 03215771 2023-10-17

and probed with human recombinant ACE-2 fused to a mouse-Fc tag (hACE2-mFc), and a secondary fluorescence-tagged anti-mouse antibody to detect ACE-2 binding to ectopically expressed vaccine candidates on the cell surface of HEK293T cells. Assuming comparable ACE-2 affinities across vaccine encoded SARS-CoV-2 variant S protein constructs, the median fluorescence intensity (MFI) measured in FACS serves as a surrogate for variant S protein surface expression.

Table 6: Sequence information about the used vaccine constructs

Underlined amino acid alterations represent those that were introduced on top of the parental Alpha vaccine sequence.

Construct	***Description****
BNT162b2	S protein sequence from SARS-CoV-2 isolate Wuhan-Hu-1 (GenBank MN908947.3) + K986P-V987P
BNT162b2(Alpha)	ΔH69/V70-ΔY144-N501Y-A570D-D614G-P681H-T716I-S982A-D1118H + K986P-V987P
BNT162b2(Alpha+SA)	ΔH69/V70-D80A-ΔY144-D215G-R246I-K417N-E484K-N501Y-A570D-D614G-P681H-T716I-S982A-D1118H + K986P-V987P
BNT162b2(Alpha;L452R+E484Q)	ΔH69/V70-ΔY144-L452R-E484Q-N501Y-A570D-D614G-P681H-T716I-S982A-D1118H + K986P-V987P

Briefly, $0.4x10^6$ HEK293T cells were seeded 6 h prior to transfection in 12-well plates. modRNA encoding for vaccine candidates was formulated using Lipofectamine™ MessengerMAX™ (ThermoFisher Scientific) according to the manufacturers' instructions prior to transfection, or LNP-formulated modRNA encoding for vaccine candidates was used to transfect cells in triplicates at a concentration of 0.15 µg/mL modRNA. Cells were incubated for 18 h at 37°C and 5% CO_2 prior to staining. Afterwards cells were harvested and incubated with a viability dye (eBioscience™ Fixable Viability Dye eFluor™ 450, ThermoFisher Scientific) and hACE2-mFc (SinoBiological) and subsequently stained with an Alexa Fluor® 647 AffiniPure Donkey Anti-Mouse IgG (H+L) secondary antibody (Jackson ImmunoResearch) prior to fixation (Fixation

CA 03215771 2023-10-17

Buffer, BioLegend). Cells were acquired using a BD FACSCelesta 2 (BD) and data was analyzed with FlowJo V10.8 (BD).

All vaccine candidate encoded S proteins were expressed on the cell surface and were able to bind to the host cell entry receptor of SARS-CoV-2 – the angiotensin-converting enzyme (ACE-2) – as determined using hACE2-mFc as a detection reagent (**Figure 7**). Expression levels were lower for Lipofectamine™ MessengerMAX™ formulated modRNA (**Figure 7A, B**) compared to those obtained with LNP-formulated modRNA (**Figure 7C, D**), as seen in the percentage of SARS-CoV-2 S protein expressing cells and the MFI of the total transfected HEK293T population. However, vaccine candidates transfected with the same method show largely comparable S protein surface expression levels [BNT162b2, BNT162b2(Alpha+SA) (**Figure 7A, B**), and BNT162b2, BNT162b2(Alpha), BNT162b2(Alpha;L452R+E484Q) (**Figure 7C, D**), respectively].

Example 3: Immunogenicity study in mice using multiple variant mutations based on the B.1.1.7 (Alpha) backbone

Three groups of 5 female BALB/c mice were injected once on Day 0 with 1 µg/animal BNT162b2(Alpha), BNT162b2(Alpha+SA), or with saline alone. Vaccine sequences are based on the B.1.1.7 (Alpha) strain adding key mutations from the SARS-CoV-2 variant B.1.351 (Beta, here abbreviated as SA), details are given in **Table 7**. Intramuscular (i.m.) injections were given in a 20 µL dose volume. Blood was collected to generate serum samples from all individual mice on Days 7, 14, 21, and 28. For pre-bleed (Pre) before immunization, 5 randomly picked mice were bled.

Table 7: Sequence information about the used vaccine constructs

Underlined amino acid alterations represent those that were introduced on top of the parental Alpha vaccine sequence.

Construct	***Description****
BNT162b2(Alpha)	ΔH69/V70-ΔY144-N501Y-A570D-D614G-P681H-T716I-S982A-D1118H + K986P-V987P

CA 03215771 2023-10-17

BNT162b2(Alpha+SA)	ΔH69/V70-D80A-ΔY144-D215G-R246I-K417N-E484K-N501Y-A570D-D614G-P681H-T716I-S982A-D1118H + K986P-V987P

Enzyme-linked immunosorbent assay (ELISA) was performed for all timepoints using either proteins including mutations derived from the B.1.1.7 or B.1.351 variants. At the end of the study on Day 28, collected serum samples were tested for neutralizing antibody responses using a VSV-SARS-CoV-2-based pseudovirus neutralization test.

For ELISA, serum samples were tested in 96-well plates to assess S or RBD protein-specific antibody concentration using the two recombinant proteins derived from either the B.1.1.7 or B.1.351 viral strain (Sino Biological; cat: 40591-V08H12 [S1(B.1.1.7)]; Sino Biological; cat: 40592-V08H86 [RBD(B.1.351)]). Briefly, each well of a MaxiSorp plate (Thermo Fisher) was coated with 100 ng recombinant protein or isotype controls. Plates were incubated over night at 4°C. After incubation, plates were washed three times with PBS + 0.01% Tween 20 (washing step) and blocked with blocking buffer for 1 h at 37°C. After a washing step, the samples or antibody controls were added, and plates were again incubated for 1 h at 37°C. Before the secondary antibody was added to the wells, another washing step was performed. Secondary antibodies conjugated with horse-reddish-peroxidase (HRP) were incubated for 45 min at 37°C. After a final washing step, TMB ONE (Biotrend Chemikalien GmbH) substrate was added to the wells and incubated for 8 min at room temperature (RT). If HRP-conjugated antibodies were present in the well, a color exchange from clear to blue was observed and reaction was stopped by adding 25% sulfuric acid (blue to yellow). Absorbance of plates was measured using an Epoch microplate reader (450 nm, reference 620 nm; BioTek).

All tested constructs induced IgG antibodies specific against S1(B.1.1.7) (**Figure 8**) and RBD(B.1.351) (**Figure 9**) at all different timepoints with titers peaking at 21 days after immunization. Results demonstrate that the inclusion of multiple variant mutations in one S protein sequence (here: Alpha backbone) in principle is a suitable antigen for vaccination. As BNT162b2(Alpha) is the construct encoding the antigen closest to the evaluated S1(B.1.1.7) recombinant protein and used as a backbone vaccine construct, it was of high relevance to understand whether immunization with the hybrid construct BNT162b2(Alpha+SA) results in

CA 03215771 2023-10-17

diminished antibody titers against S1(B.1.1.7). Looking at the single data points per timepoint testing for anti-S1(B.1.1.7) IgG antibodies(**Figure 8**), both vaccine candidates induced elevated titers compared to the buffer control with BNT162b2(Alpha+SA) inducing a weaker response against the B.1.1.7-derived S1 protein. In comparison to immunization with BNT162b2(Alpha), BNT162b2(Alpha+SA) induced either similar (at day 7) or lower B.1.1.7 S1-specific IgG titers. Focusing on significant differences between the immunized groups (**Table 8**), the IgG titers elicited by BNT162b2(Alpha+SA) compared to BNT162b2(Alpha) were significantly lower at study days 14 and 28. The result shows that even though multiple non-Alpha mutations were included, the hybrid vaccine construct is still a good vaccine candidate eliciting binding-antibody titers against the B.1.1.7 (Alpha) backbone viral protein.

Table 8: Summary of statistical significance# of anti-S1(B.1.1.7) IgG antibody titers after one immunization of Balb/c mice with different mRNA vaccine candidates encoding a P2 S construct variant.

One way ANOVA using Tukey's multiple comparisons test	**Day 7**	**Day 14**	**Day 21**	**Day 28**
buffer vs. BNT162b2(Alpha+SA)	****	***	****	****
buffer vs. BNT162b2(Alpha)	****	****	****	****
BNT162b2(Alpha+SA) vs. BNT162b2(Alpha)	ns	**	ns	****

significance is given as ****p<0.0001, ***p≤0.001, **p≤0.01, *p≤0.05, ns = not significant

Looking at the single data points per timepoint testing for anti-RBD(B.1.351) IgG antibodies (**Figure 9**), all vaccine candidates induced considerably elevated titers compared to the buffer control. With the BNT162b2(Alpha+SA) vaccine candidate being the closest to the recombinant B.1.351 variant RBD in this assay and inducing the highest response against the antigen (except day 28), results demonstrate that inclusion of multiple mutations drives the immune response towards corresponding variants. In comparison to immunization with BNT162b2(Alpha), BNT162b2(Alpha+SA) induced either similar or higher IgG titers. Focusing on significant differences between the immunized groups (**Table 9**), titers elicited by BNT162b2(Alpha+SA) compared to BNT162b2(Alpha) were significantly higher at study days 7

CA 03215771 2023-10-17

and 21, showing a clear beneficial impact when variant specific mutations were included to the B.1.1.7 (Alpha) backbone. Results show that the hybrid vaccine construct is a good vaccine candidate and that the implemented variant-specific mutations towards an additional viral strain (here the B.1.351 specific mutations) lead to an immunological benefit.

Table 9: Summary of statistical significance# of anti-RBD(B.1.351) IgG antibody titers after one immunization of Balb/C mice with different mRNA vaccine candidates encoding a P2 S construct variant.

One way ANOVA using Tukey's multiple comparisons test	Day 7	Day 14	Day 21	Day 28
buffer vs. BNT162b2(Alpha+SA)	****	****	****	***
buffer vs. BNT162b2(Alpha)	*	****	**	****
BNT162b2(Alpha+SA) vs. BNT162b2(Alpha)	****	ns	*	*

significance is given as ****p<0.0001, ***p≤0.001, **p≤0.01, *p≤0.05, ns = not significant

Analyzing the neutralizing functionality of the induced antibodies after immunization, all sera from blood samples collected 28 days after immunization were tested using the VSV/SARS-CoV-2-based pseudovirus neutralization test (pVNT). For the pVNT assay, mouse serum samples were serially diluted in duplicates in a 96-well V-bottom plate and incubated with a defined number of VSV/SARS-CoV-2 pseudovirus particles; the S protein sequences used for pseudotyping were derived from the ancestral SARS-CoV-2 Wuhan strain (Wuhan), the B.1.1.7 (Alpha) variant (mutations: Δ69/70, Δ144, N501Y, A570D, D614G, P681H, T716I, S982A, D1118H) and the B.1.351 (Beta) variant (mutations: L18F, D80A, D215G, Δ242–244, R246I, K417N, E484K, N501Y, D614G, A701V). After incubation, to allow antibodies to bind to the pseudovirus, pseudovirus/serum dilution mix was added to Vero-76 cells which were previously seeded in 96-well flat-bottom plates. Plates were incubated for 16 to 24 hours at 37°C and 5% CO_2. Vero-76 cells incubated with pseudovirus in the absence of mouse sera were used as positive controls. Vero-76 cells incubated without pseudovirus were used as negative controls. After incubation, supernatants were removed and the cells were lysed with luciferase reagent (Promega). Luminescence was recorded on a CLARIOstar® Plus microplate

CA 03215771 2023-10-17

reader (BMG Labtech), and neutralization titers were calculated as the reciprocal of the highest serum dilution that still resulted in 50% reduction in luminescence. If no neutralization was observed, an arbitrary titer value of half of the limit of detection (LOD) was reported.

All mice immunized with BNT162b2(Alpha) and BNT162b2(Alpha+SA) showed a detectable neutralizing antibody titer against Wuhan pseudovirus. However, group geometric mean 50% pseudovirus neutralizing antibody (pVN_{50}) titers were rather low, ranging from 48 for BNT162b2(Alpha) immune sera to 24 for BNT162b2(Alpha+SA) immune sera (**Figure 10A**). Compared to the Wuhan titers considerably higher pVN_{50} titers were observed against the Alpha pseudovirus, especially in serum samples drawn from animals immunized with BNT162b2(Alpha) (**Figure 10B**; geometric mean titer of 182) and to a lesser extent for BNT162b2(Alpha+SA) (geometric mean titer of 43). Higher Alpha compared to Wuhan pseudovirus titers are in line with the fact that both tested vaccine candidates encode an Alpha backbone-derived Spike protein as antigen. The lower Alpha pseudovirus neutralizing titer elicited by BNT162b2(Alpha+SA) compared to BNT162b2(Alpha) may be due to the fact that the hybrid vaccine harbors additional 'non-Alpha' mutations in the RBD [see **Table 7**; K417N and E484K for BNT162b2(Alpha+SA)], which potentially trigger a less effective neutralizing antibody response against the B.1.1.7/Alpha pseudovirus. Most importantly, on the other hand the BNT162b2(Alpha+SA) vaccine induced considerably higher pVN_{50} titers against the Beta pseudovirus when compared to the BNT162b2(Alpha) vaccine (**Figure 10C**; group geometric mean titer of 96 vs. 18), showing that the implemented Beta variant specific mutations led to a more Beta-specific antibody response

In summary, using multiple variant mutations in a SARS-CoV-2 P2 S vaccine construct backbone slightly alters the immune response towards the viral antigen corresponding to the amino acid changes used in the backbone.

Example 4: Immunogenicity study in mice including minimal variant mutations based on the B.1.1.7 (Alpha) background

Three groups of 5 female BALB/c mice were injected once on Day 0 with 1 µg/animal of BNT162b2(Alpha), BNT162b2(Alpha;L452R+E484Q), or with saline alone. For sequence specification of the constructs see **Table 10**. Intramuscular (i.m.) injections were given in a 20

CA 03215771 2023-10-17

µL dose volume. Blood was collected to generate serum samples from all individual mice on Day 28.

Table 10: Sequence information about the used vaccine constructs

Construct	*Description**
BNT162b2(Alpha)	ΔH69/V70-ΔY144-N501Y-A570D-D614G-P681H-T716I-S982A-D1118H + K986P-V987P
BNT162b2(Alpha;L452R+E484Q)	ΔH69/V70-ΔY144- L452R- E484Q-N501Y-A570D-D614G-P681H-T716I-S982A-D1118H + K986P-V987P

* given mutation in reference to the BNT162b construct derived from SARS-CoV-2 Wuhan-Hu-1 (GenBank MN908947.3) + K986P-V987P

To understand the virus neutralizing antibody response elicited by the respective immunization, serum samples were tested using the VSV/SARS-CoV-2-based pseudovirus neutralization test (pVNT). For the pVNT assay, mouse serum samples were serially diluted in duplicates in a 96-well V-bottom plate and incubated with a defined number of VSV/SARS-CoV-2 pseudovirus particles; the S protein sequences used for pseudotyping were derived from the SARS-CoV-2 B.1.1.7 (Alpha) variant (mutations: Δ69/70, Δ144, N501Y, A570D, D614G, P681H, T716I, S982A, D1118H), the B.1.617.1 (Kappa) variant (mutations: L452R, E484Q, D614G, P681R), or the B.1.617.2 (Delta) variant (mutations: T19R, G142D, E156G, Δ157/158, K417N, L452R, T478K, D614G, P681R, D950N). After incubation, to allow antibodies to bind to the pseudovirus, pseudovirus/serum dilution mix was added to Vero-76 cells which were previously seeded in 96-well flat-bottom plates. Plates were incubated for 16 to 24 hours at 37°C and 5% CO_2. Vero-76 cells incubated with pseudovirus in the absence of mouse sera were used as positive controls. Vero-76 cells incubated without pseudovirus were used as negative controls. After incubation, supernatants were removed and the cells were lysed with luciferase reagent (Promega). Luminescence was recorded on a CLARIOstar® Plus microplate reader (BMG Labtech), and neutralization titers were calculated as the reciprocal of the highest serum dilution that still resulted in 50% reduction in luminescence. If no neutralization was observed, an arbitrary titer value of half of the limit of detection (LOD) was reported.

CA 03215771 2023-10-17

Immunization of mice with BNT162b2(Alpha) or BNT162b2(Alpha;L452R+E484Q) induced largely comparable neutralizing antibody titers against the Alpha pseudovirus (geometric mean pVN50 titer of 291 and 221, respectively). However, of note, the BNT162b2(Alpha;L452R+E484Q) vaccine candidate, encoding the Alpha Spike antigen backbone with the additional Kappa and Delta variant mutations L452R and E484Q, trended to induce higher neutralizing activity against Kappa (geometric mean titer of 192 vs. 110) and Delta pseudovirus (geometric mean titer of 127 vs. 84) when compared to the BNT162b2(Alpha) vaccine (**Figure 11**).

In addition, binding capacity of antibodies generated after vaccination with the two candidates was analyzed performing a multiplex analysis. Briefly, a multiplex assay was performed according to the provider's general protocol of COVID-19 serology mouse kits which employs a sandwich immunoassay technique (Meso Scale Diagnostics, LLC). Binding mouse antibodies can be detected with a "Sulfo-Tag" conjugated secondary antibody. The multiplex reader instrument MESO QuickPlex SQ 120 (Meso Scale Diagnostics, LLC) is used in a final step that measures the light emitted from the Sulfo-Tag. For the analysis, the Multiplex assay included recombinant proteins as listed in **Table 11.**

Table 11: Recombinant proteins included in the assay

Recombinant protein	***Description****
SARS-CoV-2 S **(B.1.1.529; BA.1 (Omicron sub-lineage))**	A67V, ΔH69-V70, T95I, G142D, Δ143-145, Δ211/L212I, ins214EPE, G339D, S371L, S373P, S375F, K417N, N440K, G446S, S477N, T478K, E484A, Q493R, G496S, Q498R, N501Y, Y505H, T547K, D614G, H655Y, N679K, P681H, N764K, D796Y, N856K, Q954H, N969K, L981F
SARS-CoV-2 S **(BA.1+L452R)**	A67V, Δ69-70, T95I, G142D/Δ143-145, Δ211/L212I, ins214EPE, G339D, S371L, S373P, S375F, K417N, N440K, L452R, G446S, S477N, T478K, E484A, Q493R, G496S, Q498R, N501Y, Y505H, T547K, D614G, H655Y, N679K, P681H, N764K, D796Y, N856K, Q954H, N969K, L981F

CA 03215771 2023-10-17

SARS-CoV-2 S **(B.1.1.7)**	ΔH69-V70, ΔY144, N501Y, A570D, D614G, P681H, T716I, S982A, D1118H

* given mutation in reference to the BNT162b construct derived from SARS-CoV-2 Wuhan-Hu-1 (GenBank MN908947.3) + K986P-V987P

While SARS-CoV-2 S (B.1.1.7) is mirroring a homologue testing system, meaning that the vaccine candidate backbone BNT162b2(Alpha) is identical to the used recombinant protein target used for read-out, testing with the SARS-CoV-2 S (BA.1) recombinant protein represents a heterologous testing system. Likewise, the SARS-CoV-2 S (BA.1+L452R) recombinant protein represents a heterologous testing system, however shares one more identical amino acid to the BNT162b2(Alpha;L452R+E484Q) vaccine candidate.

Electrochemiluminescent (ECL) signals were analyzed and all animals immunized with either BNT162b2(Alpha) or BNT162b2(Alpha;L452R+E484Q) developed a strong antibody binding titer that was significantly higher compared to buffer control animals (**Figure 12, Table 12;** group mean comparison using One-way ANOVA; Tukey's multiple comparison test). Immunization with BNT162b2(Alpha;L452R+E484Q) led to higher titers against all tested antigens, however the inclusion of the L452R+E484Q mutations induced a significant higher antibody binding to the SARS-CoV-2 S (BA.1+L452R) antigen which shares one identical mutation relevant for antibody binding. Of note, overall the binding of antibodies against the BA.1 S protein variant was strongly reduced compared to the B.1.1.7 S protein, which serves as backbone vaccine antigen.

Table 12: Summary of statistical significance# of anti-S IgG antibody titers after one immunization in Balb/C mice.

One way ANOVA using Tukey's multiple comparisons test	**SARS-CoV-2 S (B.1.1.529; BA.1**	**SARS-CoV-2 S (BA.1+L452R)**	**SARS-CoV-2 S (B.1.1.7)**
buffer vs. BNT162b2 (Alpha)	**	***	****
buffer vs. BNT162b2 (Alpha;L452R+E484Q)	***	****	****

CA 03215771 2023-10-17

BNT162b2 (Alpha) vs. BNT162b2 (Alpha;L452R+E484Q)	ns	*	ns

significance is given as ****p<0.0001, ***p≤0.001, **p≤0.01, *p≤0.05, ns = not significant

In summary results demonstrated that the inclusion of a mutation occurring in viral subtypes into a backbone construct leads to an immunogenic vaccine construct comparable to the corresponding viral backbone. Testing antigens including mutations towards the included amino acids showed a slight benefit in antibody binding arguing for the feasibility of this directed antigen design approach.

Claims

1. A method comprising the steps:

a) identifying amino acid positions in a parental SARS-CoV-2 spike protein (S protein) which are modified compared to the corresponding amino acid positions of one or more SARS-CoV-2 S protein variants; and

b) providing an amino acid sequence comprising at least a fragment of the parental SARS-CoV-2 S protein, wherein amino acid positions in the at least a fragment of a parental SARS-CoV-2 S protein are modified to comprise amino acids found in the corresponding amino acid positions of one or more SARS-CoV-2 S protein variants, or a nucleotide sequence encoding the modified amino acid sequence.

2. The method of claim 1, which comprises repeating step b) to provide two or more of the modified amino acid sequences, or two or more of the nucleotide sequences encoding two or more of the modified amino acid sequences.

3. The method of claim 2, wherein the two or more modified amino acid sequences are based on the same parental SARS-CoV-2 S protein.

4. The method of claim 2 or 3, wherein the amino acid modifications in the two or more modified amino acid sequences are at least partially different.

5. The method of any one of claims 1 to 4, wherein providing the nucleotide sequence comprises:

b') substituting codons of a nucleotide sequence encoding an amino acid sequence comprising at least a fragment of a parental SARS-CoV-2 S protein with other codons to obtain a mutated nucleotide sequence that encodes a modified amino acid sequence, wherein amino acid positions in the at least a fragment of a parental SARS-CoV-2 S protein are modified to comprise amino acids found in the corresponding amino acid positions of one or more SARS-CoV-2 S protein variants.

CA 03215771 2023-10-17

6. The method of claim 5, which comprises repeating step b') to provide two or more of the mutated nucleotide sequences encoding two or more of the modified amino acid sequences.

7. The method of claim 6, wherein the two or more modified amino acid sequences are based on the same parental SARS-CoV-2 S protein.

8. The method of claim 6 or 7, wherein the amino acid modifications in the two or more modified amino acid sequences are at least partially different.

9. A method comprising the steps:
a) identifying amino acid positions in a parental SARS-CoV-2 spike protein (S protein) which are modified compared to the corresponding amino acid positions of one or more SARS-CoV-2 S protein variants;
b) substituting codons of a first nucleotide sequence encoding an amino acid sequence comprising at least a fragment of a parental SARS-CoV-2 S protein with other codons to obtain a first mutated nucleotide sequence that encodes a modified amino acid sequence, wherein amino acid positions in the at least a fragment of a parental SARS-CoV-2 S protein are modified to comprise amino acids found in the corresponding amino acid positions of one or more SARS-CoV-2 S protein variants; and
c) substituting codons of a second nucleotide sequence encoding an amino acid sequence comprising at least a fragment of a parental SARS-CoV-2 S protein with other codons to obtain a second mutated nucleotide sequence that encodes a modified amino acid sequence, wherein amino acid positions in the at least a fragment of a parental SARS-CoV-2 S protein are modified to comprise amino acids found in the corresponding amino acid positions of one or more SARS-CoV-2 S protein variants,
wherein the amino acid modifications in b) at least partially differ from the amino acid modifications in c).

10. The method of claim 9, wherein the at least a fragment of a parental SARS-CoV-2 S protein comprised in the amino acid sequence encoded by the first nucleotide sequence and the at least a fragment of a parental SARS-CoV-2 S protein comprised in the amino acid sequence encoded by the second nucleotide sequence are identical.

CA 03215771 2023-10-17

11. The method of claim 9 or 10, wherein the amino acid sequence encoded by the first nucleotide sequence and the amino acid sequence encoded by the second nucleotide sequence are identical.

12. The method of any one of claims 9 to 11, wherein the first nucleotide sequence and the second nucleotide sequence are identical.

13. The method of any one of claims 9 to 12, wherein one or more of the modified amino acid positions in the modified amino acid sequence encoded by the first mutated nucleotide sequence differ from the modified amino acid positions in the modified amino acid sequence encoded by the second mutated nucleotide sequence.

14. The method of any one of claims 9 to 13, wherein one or more amino acids in modified amino acid positions modified in the modified amino acid sequence encoded by the first mutated nucleotide sequence and in the modified amino acid sequence encoded by the second mutated nucleotide sequence differ from each other.

15. The method of any one of claims 1 to 14, wherein the amino acid sequence comprising at least a fragment of a parental SARS-CoV-2 S protein comprises the amino acid sequence of the N-terminal domain (NTD) and/or receptor binding domain (RBD) of a SARS-CoV-2 S protein.

16. The method of any one of claims 1 to 15, wherein the amino acid sequence comprising at least a fragment of a parental SARS-CoV-2 S protein comprises the amino acid sequence of a full-length SARS-CoV-2 S protein.

17. The method of any one of claims 1 to 16, further comprising providing nucleic acid comprising the nucleotide sequence encoding a modified amino acid sequence.

18. The method of any one of claims 1 to 17, further comprising providing a vaccine comprising nucleic acid comprising the nucleotide sequence encoding a modified amino acid sequence.

CA 03215771 2023-10-17

19. The method of claim 17 or 18, wherein the nucleic acid is RNA.

20. The method of any one of claims 1 to 19, which is a method of generating a SARS-CoV-2 vaccine.

21. The method of any one of claims 18 to 20, wherein the vaccine is an RNA vaccine.

22. The method of any one of claims 18 to 21, wherein the vaccine has a reduced risk for immune escape.

23. The method of any one of claims 1 to 22, wherein one or more of the modified amino acid positions are located within the N-terminal domain (NTD) and/or the receptor binding domain (RBD) of a SARS-CoV-2 S protein.

24. The method of any one of claims 1 to 23, wherein the modified amino acid positions are amino acid positions at which the amino acid sequence of the one or more SARS-CoV-2 S protein variants differs from the amino acid sequence of the parental SARS-CoV-2 S protein.

25. The method of any one of claims 1 to 24, wherein the modified amino acid positions are amino acid positions at which the amino acid sequence of the one or more SARS-CoV-2 S protein variants differs from the amino acid sequence of wildtype SARS-CoV-2 S protein.

26. The method of any one of claims 1 to 25, wherein the modified amino acid positions are potential sites for escape mutants of SARS-CoV-2.

27. The method of claim 26, wherein the escape mutants of SARS-CoV-2 are antibody escape mutants of SARS-CoV-2.

28. The method of claim 26 or 27, wherein the escape mutants of SARS-CoV-2 are resistant to neutralization by antibody against SARS-CoV-2 S protein.

29. The method of any one of claims 26 to 28, wherein the SARS-CoV-2 S protein of the escape mutants of SARS-CoV-2 shows reduced antibody binding.

30. The method of any one of claims 27 to 29, wherein the antibody is used for treating a patient infected with SARS-CoV-2.

31. The method of any one of claims 27 to 29, wherein the antibody is generated in a patient that has been treated with a SARS-CoV-2 vaccine.

32. The method of any one of claims 1 to 31, wherein the parental SARS-CoV-2 S protein is modified compared to wildtype SARS-CoV-2 S protein.

33. The method of any one of claims 1 to 32, wherein in the modified amino acid sequence amino acid positions in the parental SARS-CoV-2 S protein which are modified compared to wildtype SARS-CoV-2 S protein are not modified.

34. The method of any one of claims 1 to 33, wherein the parental SARS-CoV-2 S protein is the S protein of a parental SARS-CoV-2 strain.

35. The method of claim 34, wherein the parental SARS-CoV-2 strain is a natural isolate, or the parental SARS-CoV-2 strain is a mutant of a natural isolate.

36. The method of claim 34 or 35, wherein the parental SARS-CoV-2 strain is a SARS-CoV-2 variant strain that is prevalent or rapidly spreading.

37. The method of any one of claims 34 to 36, wherein the parental SARS-CoV-2 strain is a SARS-CoV-2 variant that is a variant of concern.

38. The method of any one of claims 34 to 37, wherein the parental SARS-CoV-2 strain is B.1.1.7.

CA 03215771 2023-10-17

39. The method of any one of claims 1 to 38, wherein the one or more SARS-CoV-2 S protein variants are modified compared to wildtype SARS-CoV-2 S protein.

40. The method of any one of claims 1 to 39, wherein the one or more SARS-CoV-2 S protein variants are modified compared to the parental SARS-CoV-2 S protein.

41. The method of any one of claims 1 to 40, wherein in the modified amino acid sequence amino acid positions in the one or more SARS-CoV-2 S protein variants which are modified compared to wildtype SARS-CoV-2 S protein and/or the parental SARS-CoV-2 S protein are modified.

42. The method of any one of claims 1 to 41, wherein one or more of the one or more SARS-CoV-2 S protein variants are the S proteins of one or more SARS-CoV-2 strains.

43. The method of claim 42, wherein one or more of the one or more SARS-CoV-2 strains are natural isolates, or one or more of the one or more SARS-CoV-2 strains are mutants of a natural isolate.

44. The method of claim 42 or 43, wherein one or more of the one or more SARS-CoV-2 strains are SARS-CoV-2 variant strains that are prevalent or rapidly spreading.

45. The method of any one of claims 42 to 44, wherein one or more of the one or more SARS-CoV-2 strains are SARS-CoV-2 variant strains that are variants of concern.

46. The method of any one of claims 42 to 45, wherein one or more of the one or more SARS-CoV-2 strains are selected from the group consisting of B.1.351, B.1.1.298, B.1.427/B.1.429, B.1.526, and P1.

47. The method of any one of claims 42 to 46, wherein the parental SARS-CoV-2 strain and the one or more SARS-CoV-2 strains are SARS-CoV-2 variant strains that are prevalent or rapidly spreading.

CA 03215771 2023-10-17

48. The method of any one of claims 42 to 47, wherein the parental SARS-CoV-2 strain and the one or more SARS-CoV-2 strains are SARS-CoV-2 variant strains that are variants of concern.

49. The method of any one of claims 1 to 48, wherein the parental SARS-CoV-2 S protein and the one or more SARS-CoV-2 S protein variants are modified compared to wildtype SARS-CoV-2 S protein.

50. The method of any one of claims 42 to 49, wherein the parental SARS-CoV-2 strain is B.1.1.7 and the one or more SARS-CoV-2 strains are selected from the group consisting of B.1.351, B.1.1.298, B.1.427/B.1.429, B.1.526, and P1.

51. The method of any one of claims 1 to 50, wherein the one or more SARS-CoV-2 S protein variants comprise SARS-CoV-2 S protein variants of at least two SARS-CoV-2 strains.

52. The method of any one of claims 9 to 51, wherein the one or more SARS-CoV-2 S protein variants in b) are different from the one or more SARS-CoV-2 S protein variants in c).

53. The method of any one of claims 9 to 52, wherein the one or more SARS-CoV-2 S protein variants in b) are the SARS-CoV-2 S protein variants of B.1.427/B.1.429, and B.1.526, and the one or more SARS-CoV-2 S protein variants in c) are the SARS-CoV-2 S protein variants of B.1.351, P.1 and B.1.1.298.

54. The method of any one of claims 1 to 53, wherein in the modified amino acid sequence amino acid modifications in the parental SARS-CoV-2 S protein compared to wildtype SARS-CoV-2 S protein do not interfere with amino acid modifications in the modified amino acid positions.

55. The method of any one of claims 1 to 54, wherein in the modified amino acid sequence amino acid modifications in the parental SARS-CoV-2 S protein compared to wildtype SARS-CoV-2 S protein are not in close spatial distance to amino acid modifications in the modified amino acid positions.

CA 03215771 2023-10-17

56. The method of any one of claims 1 to 55, wherein in the modified amino acid sequence modifications in the modified amino acid positions do not result in major structural rearrangements.

57. The method of any one of claims 1 to 56, wherein the amino acids in the modified amino acid positions are surface exposed.

58. The method of any one of claims 1 to 57, wherein the modified amino acid positions comprise at least two amino acid positions.

59. The method of any one of claims 9 to 58, wherein the modified amino acid positions in b) and c) each comprise at least two amino acid positions.

60. The method of any one of claims 1 to 59, wherein the modified amino acid positions comprise two or more selected from the group consisting of:
18, 20, 26, 80, 138, 144, 190, 215, 246, 253, 417, 439, 452, 453, 477, 484, 501, 570, 701, 716,

140, 345, 346, 352, 378, 406, 420, 440, 441, 444, 445, 446, 450, 455, 460, 475, 478, 485, 486, 487, 489, 490, 493, 494, 499,

142, 145, 146, 147, 150, 152, 154, 156, 157, 158, 164, 247, 248, 249, 250, 251, 252, 254, 255, 258, 365, 369, 370, 374, 376, 384, 405, 408, 415, 421, 443, 447, 448, 456, 472, 473, 476, 496, 498, 500, 504.

61. The method of any one of claims 1 to 60, wherein the modifications in the modified amino acid positions comprise two or more selected from the group consisting of:
18F, 20N, 26S, 80Y, 138Y, 144F, 190S, 215A, 246I, 253G, 417N, 439K, 452R, 453F, 477N, 484K, 501Y, 570D, 701V, 716I,

140L, 345A, 346K, 352S, 378N, 406Q, 420, 440K, 441F, 444, 445A, 446V, 450K, 455F, 460I, 475V, 478I, 485V, 486L, 487D, 489, 490S, 493L, 494P, 499H,

142S, 145H, 146Y, 147N, 150R, 152C, 154Q, 156A, 157L, 158G, 164T, 247G, 248H, 249S, 250N, 251S, 252V, 254F, 255F, 258L, 365D, 369C, 370S, 374L, 376I, 384L, 405Y, 408I, 415N, 421, 443A, 447V, 448Y, 456L, 472V, 473F, 476S, 496C, 498H, 500I, 504D.

62. The method of any one of claims 1 to 61, wherein the modifications in the modified amino acid positions comprise two or more selected from the group consisting of:
L18F, T20N, P26S, D80Y, D138Y, Y144F, R190S, D215A, R246I, D253G, K417N, N439K, L452R, Y453F, S477N, E484K, N501Y, A570D, A701V, T716I,

F140L, T345A, R346K, A352S, K378N, E406Q, D420, N440K, L441F, K444, V445A, G446V, N450K, L455F, N460I, A475V, T478I, G485V, F486L, N487D, Y489, F490S, Q493L, S494P, P499H,

G142S, Y145H, H146Y, K147N, K150R, W152C, E154Q, E156A, F157L, R158G, N164T, S247G, Y248H, L249S, T250N, P251S, G252V, S254F, S255F, W258L, Y365D, Y369C, N370S, F374L, T376I, P384L, D405Y, R408I, T415N, Y421, S443A, G447V, N448Y, F456L, I472V, Y473F, G476S, G496C, Q498H, T500I, G504D.

63. A method comprising the steps:
a) providing a nucleic acid comprising a first nucleotide sequence encoding an amino acid sequence comprising at least a fragment of a parental SARS-CoV-2 S protein, wherein amino acid positions in the at least a fragment of a parental SARS-CoV-2 S protein are modified to comprise amino acids found in the corresponding amino acid positions of one or more SARS-CoV-2 S protein variants; and
b) providing a nucleic acid comprising a second nucleotide sequence encoding an amino acid sequence comprising at least a fragment of a parental SARS-CoV-2 S protein, wherein amino acid positions in the at least a fragment of a parental SARS-CoV-2 S protein are modified to comprise amino acids found in the corresponding amino acid positions of one or more SARS-CoV-2 S protein variants,
wherein the amino acid modifications in b) at least partially differ from the amino acid modifications in a).

CA 03215771 2023-10-17

64. The method of claim 63, wherein the nucleic acid is RNA.

65. A medical preparation comprising:
a) a nucleic acid comprising a first nucleotide sequence encoding an amino acid sequence comprising at least a fragment of a parental SARS-CoV-2 S protein, wherein amino acid positions in the at least a fragment of a parental SARS-CoV-2 S protein are modified to comprise amino acids found in the corresponding amino acid positions of one or more SARS-CoV-2 S protein variants; and
b) a nucleic acid comprising a second nucleotide sequence encoding an amino acid sequence comprising at least a fragment of a parental SARS-CoV-2 S protein, wherein amino acid positions in the at least a fragment of a parental SARS-CoV-2 S protein are modified to comprise amino acids found in the corresponding amino acid positions of one or more SARS-CoV-2 S protein variants,
wherein the amino acid modifications in b) at least partially differ from the amino acid modifications in a).

66. The medical preparation of claim 65, wherein the nucleic acid is RNA.

67. The medical preparation of claim 66, wherein the RNA is formulated in lipid nanoparticles (LNP).

68. The medical preparation of any one of claims 65 to 67, which is a pharmaceutical composition.

69. The medical preparation of any one of claims 65 to 68, which is a vaccine.

70. The medical preparation of any one of claims 65 to 69, which is a kit.

71. The medical preparation of claim 70, further comprising instructions for use of the medical preparation for vaccination against SARS-CoV-2 infection.

72. The medical preparation of any one of claims 65 to 71 for pharmaceutical use.

CA 03215771 2023-10-17

73. The medical preparation of claim 72, wherein the pharmaceutical use comprises vaccination against SARS-CoV-2 infection.

74. A method of inducing an immune response against SARS-CoV-2 in a subject comprising administering to the subject the medical preparation of any one claims 65 to 73.

75. The method of claim 74, which is a method for prophylactic treatment against SARS-CoV-2 infection.

76. The method of claim 74 or 75, which is a method for vaccination against SARS-CoV-2 infection.

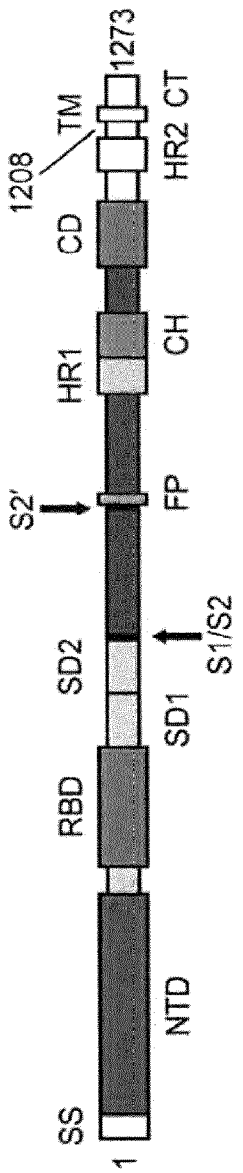


Figure 1

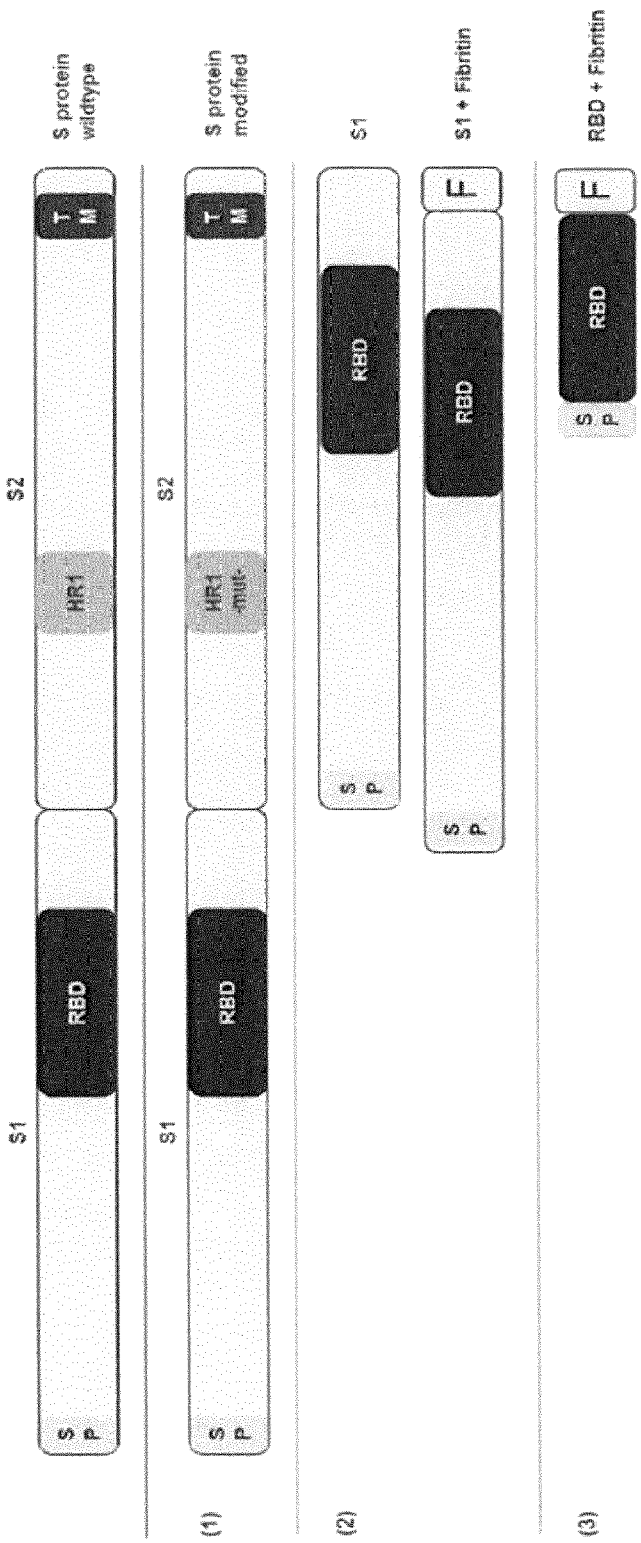

S protein wildtype
S protein modified
S1
S1 + Fibritin
RBD + Fibritin
S2
S1
RBD
HR1
HR1 -mut-
T M
F
S P
(1)
(2)
(3)

Figure 2

Figure 3

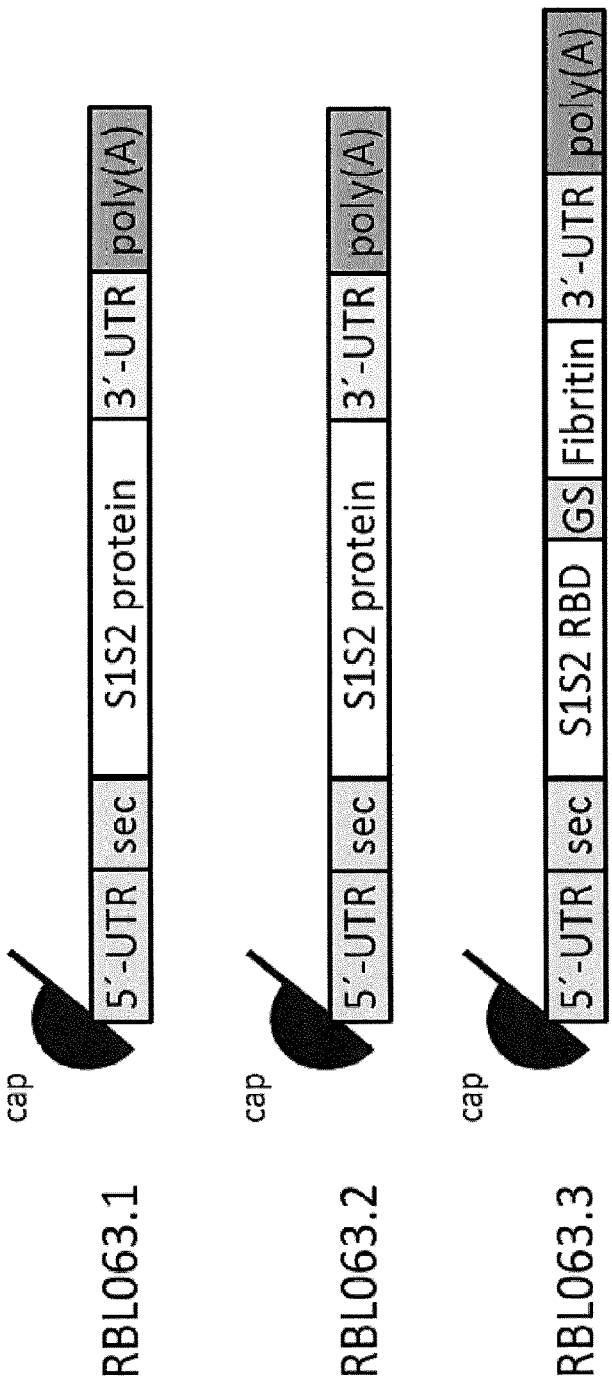

cap
5´-UTR
sec
S1S2 protein
3´-UTR
poly(A)
RBL063.1
cap
5´-UTR
sec
S1S2 protein
3´-UTR
poly(A)
RBL063.2
cap
5´-UTR
sec
S1S2 RBD
GS
Fibritin
3´-UTR
poly(A)
RBL063.3

Figure 4

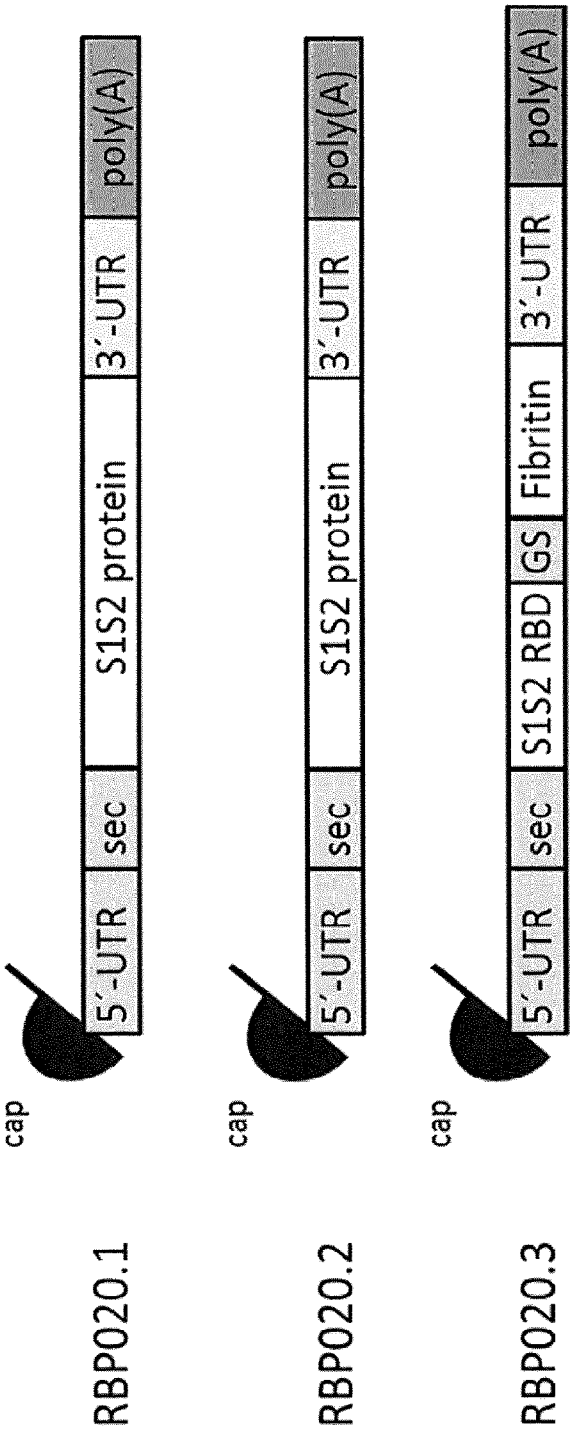
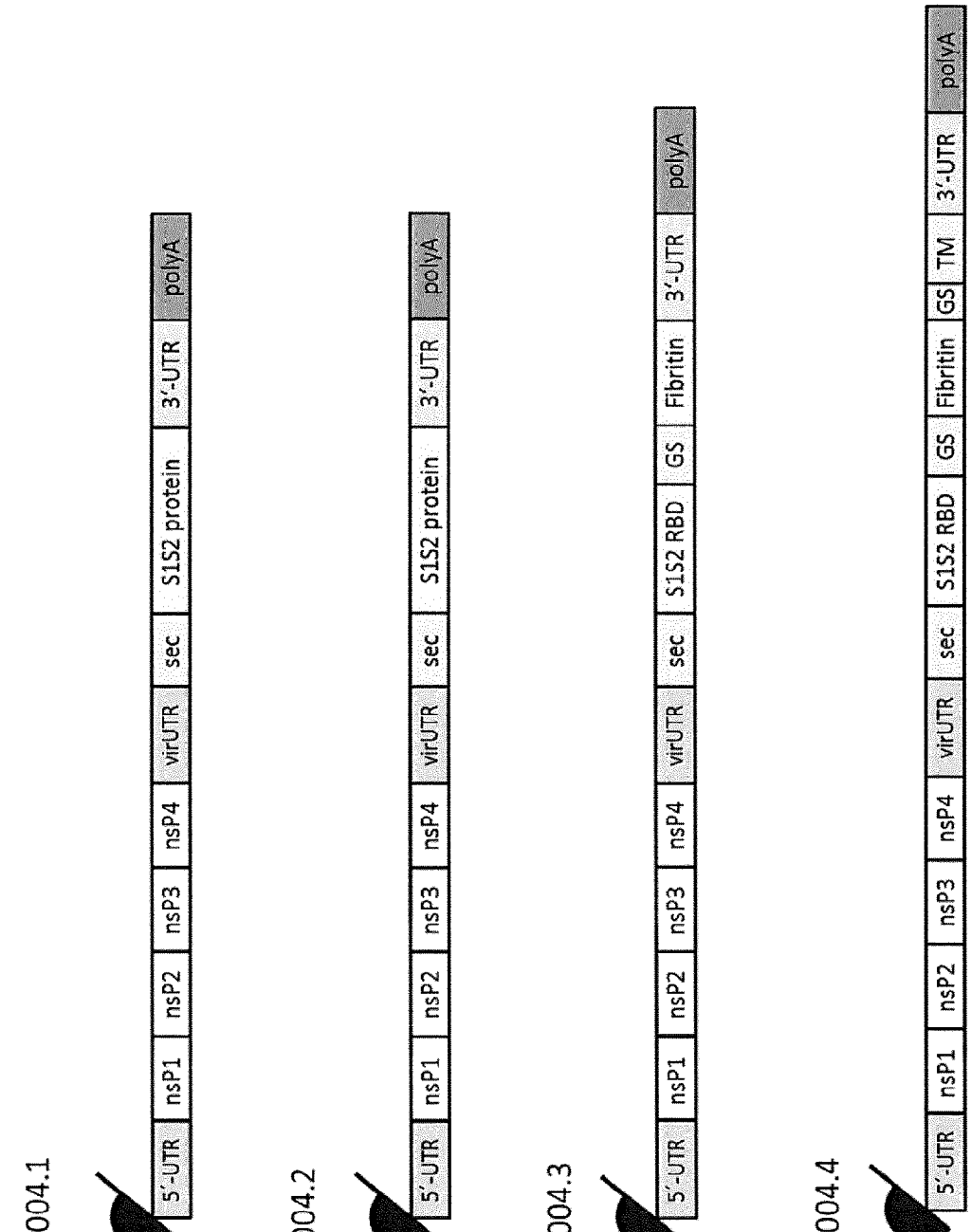
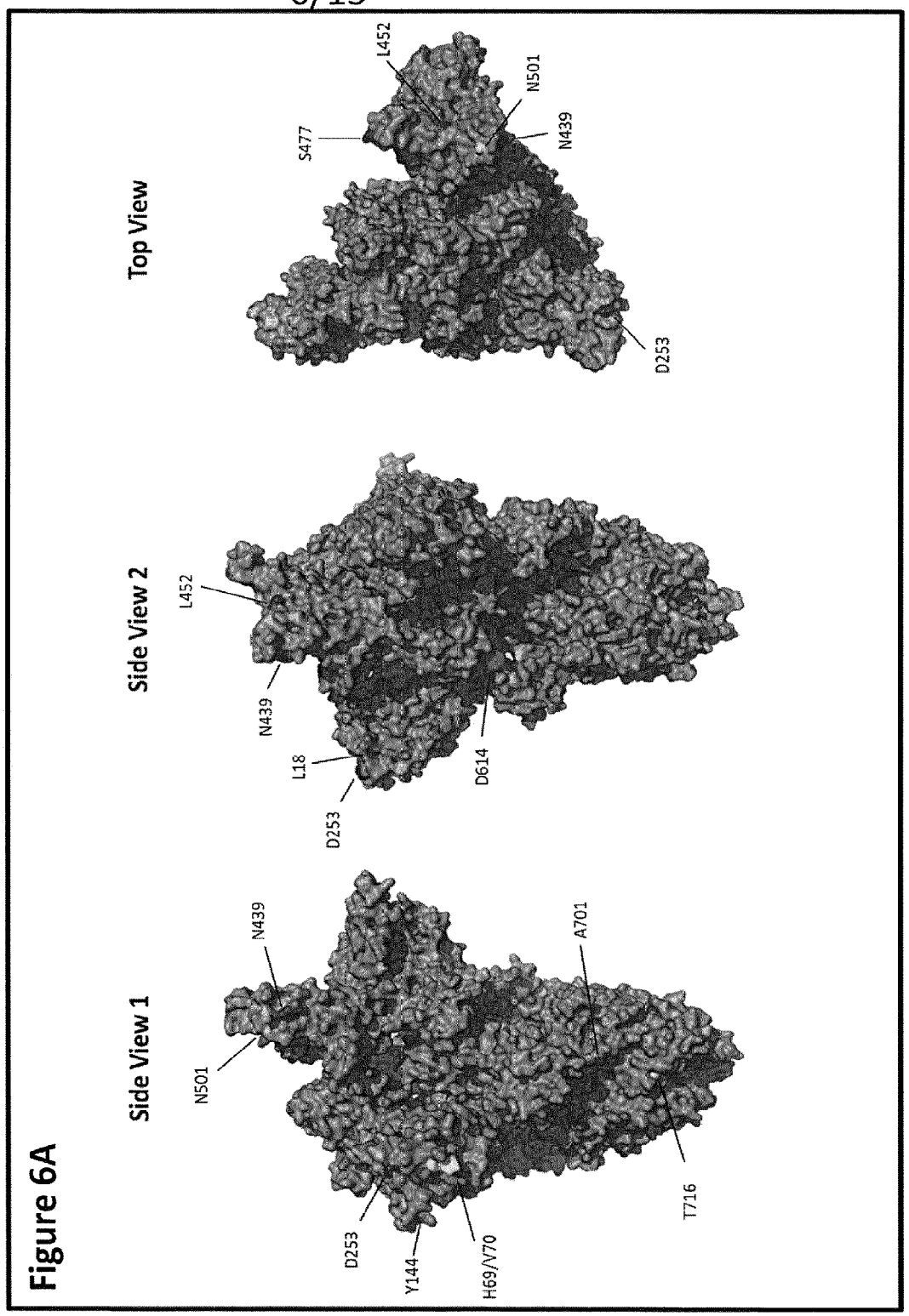
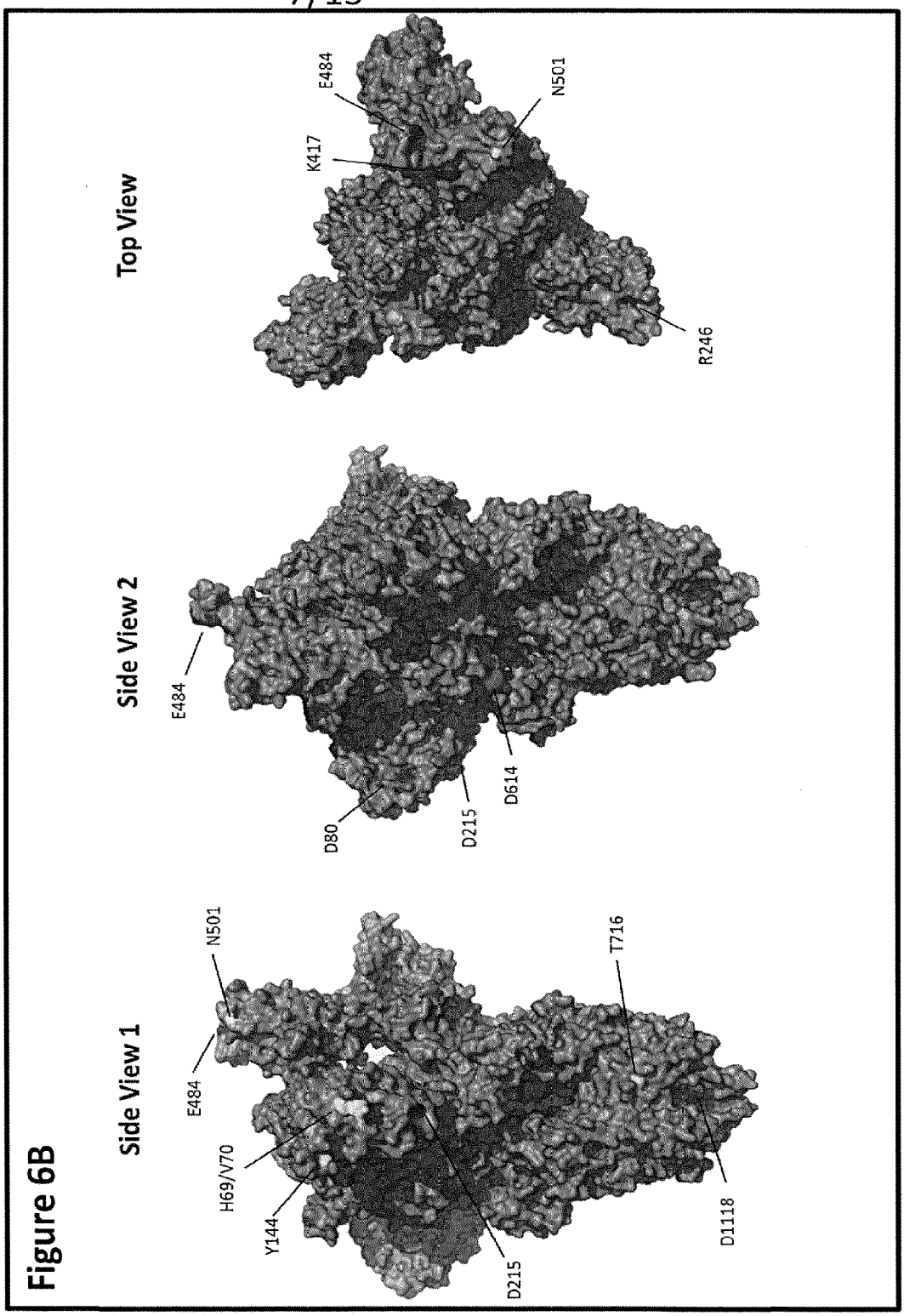

cap
5´-UTR
sec
S1S2 protein
3´-UTR
poly(A)
RBP020.1
cap
5´-UTR
sec
S1S2 protein
3´-UTR
poly(A)
RBP020.2
cap
5´-UTR
sec
S1S2 RBD
GS
Fibritin
3´-UTR
poly(A)
RBP020.3

Figure 5


RBS004.1
cap
5'-UTR
nsP1
nsP2
nsP3
nsP4
virUTR
sec
S1S2 protein
3'-UTR
polyA
RBS004.2
cap
5'-UTR
nsP1
nsP2
nsP3
nsP4
virUTR
sec
S1S2 protein
3'-UTR
polyA
RBS004.3
cap
5'-UTR
nsP1
nsP2
nsP3
nsP4
virUTR
sec
S1S2 RBD
GS
Fibritin
3'-UTR
polyA
RBS004.4
cap
5'-UTR
nsP1
nsP2
nsP3
nsP4
virUTR
sec
S1S2 RBD
GS
Fibritin
GS
TM
3'-UTR
polyA


Figure 6A
Side View 1
N501
N439
D253
Y144
H69/V70
A701
T716
Side View 2
L452
N439
L18
D253
D614
Top View
S477
L452
N501
N439
D253


Figure 6B
Side View 1
N501
E484
H69/V70
Y144
D215
T716
D1118
Side View 2
E484
D80
D215
D614
Top View
E484
N501
K417
R246

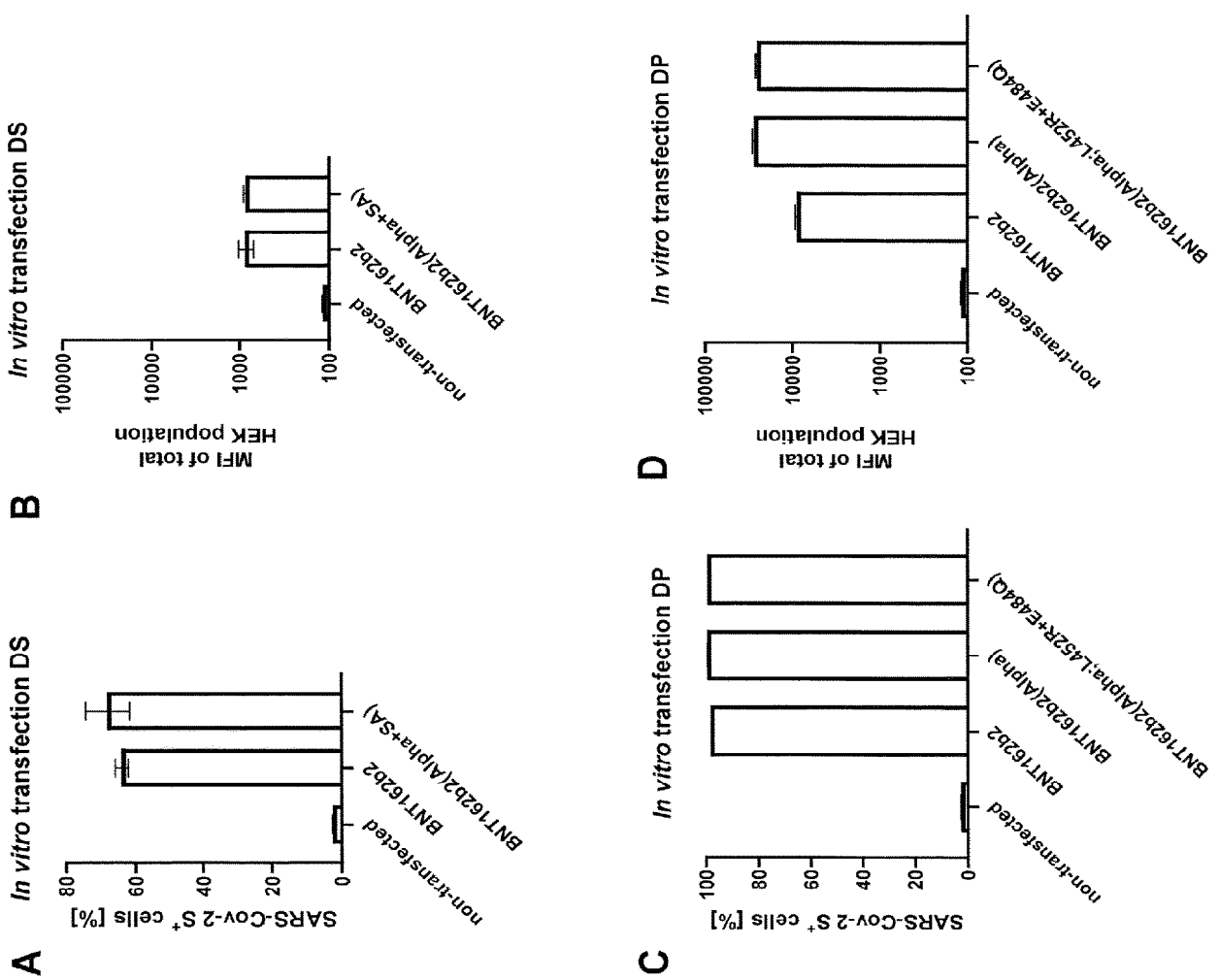

A
In vitro transfection DS
SARS-CoV-2 S+ cells [%]
0
20
40
60
80
non-transfected
BNT162b2
BNT162b2(Alpha+SA)
B
In vitro transfection DS
MFI of total HEK population
100
1000
10000
100000
non-transfected
BNT162b2
BNT162b2(Alpha+SA)
C
In vitro transfection DP
SARS-CoV-2 S+ cells [%]
0
20
40
60
80
100
non-transfected
BNT162b2
BNT162b2(Alpha)
BNT162b2(Alpha;L452R+E484Q)
D
In vitro transfection DP
MFI of total HEK population
100
1000
10000
100000
non-transfected
BNT162b2
BNT162b2(Alpha)
BNT162b2(Alpha;L452R+E484Q)

Figure 7

Figure 8

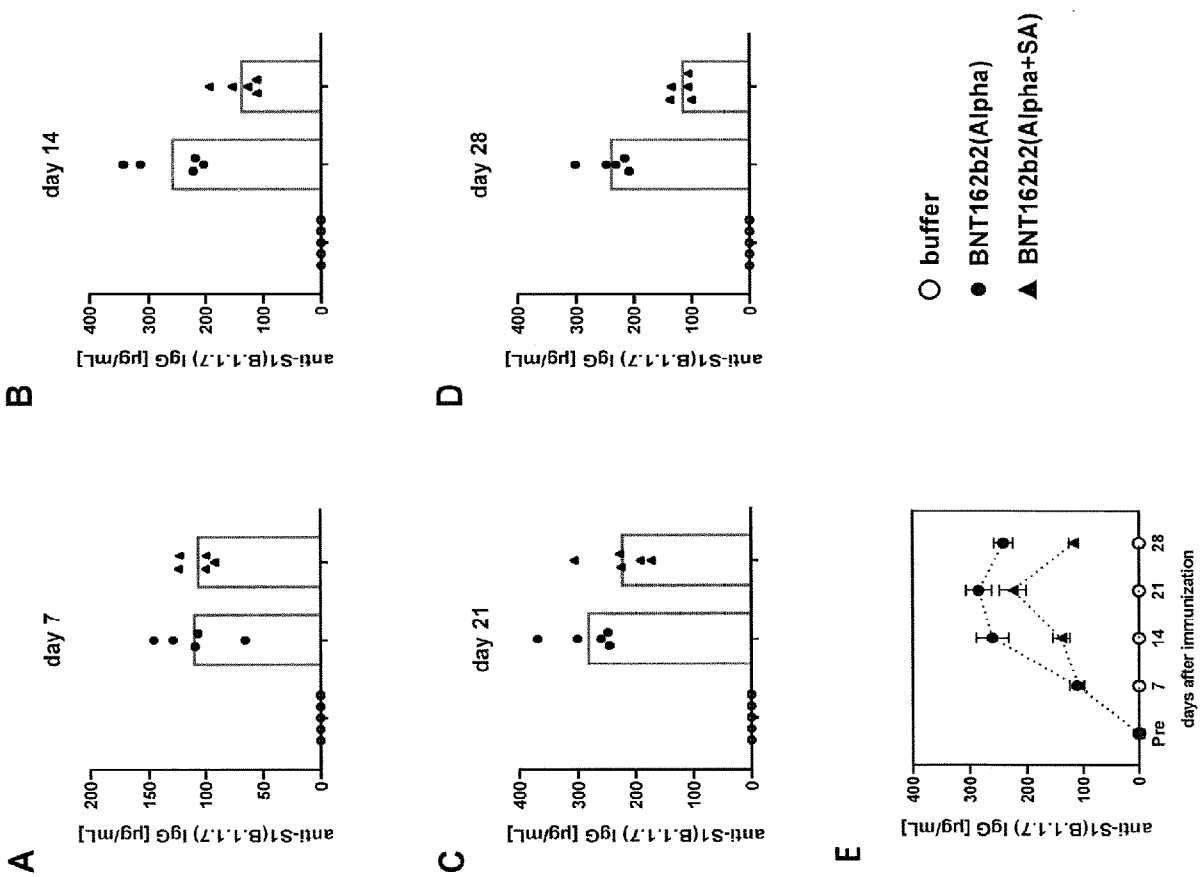

A
day 7
anti-S1(B.1.1.7) IgG [µg/mL]
0
50
100
150
200
B
day 14
anti-S1(B.1.1.7) IgG [µg/mL]
0
100
200
300
400
C
day 21
anti-S1(B.1.1.7) IgG [µg/mL]
0
100
200
300
400
D
day 28
anti-S1(B.1.1.7) IgG [µg/mL]
0
100
200
300
400
E
anti-S1(B.1.1.7) IgG [µg/mL]
0
100
200
300
400
Pre
7
14
21
28
days after immunization
buffer
BNT162b2(Alpha)
BNT162b2(Alpha+SA)

Figure 9

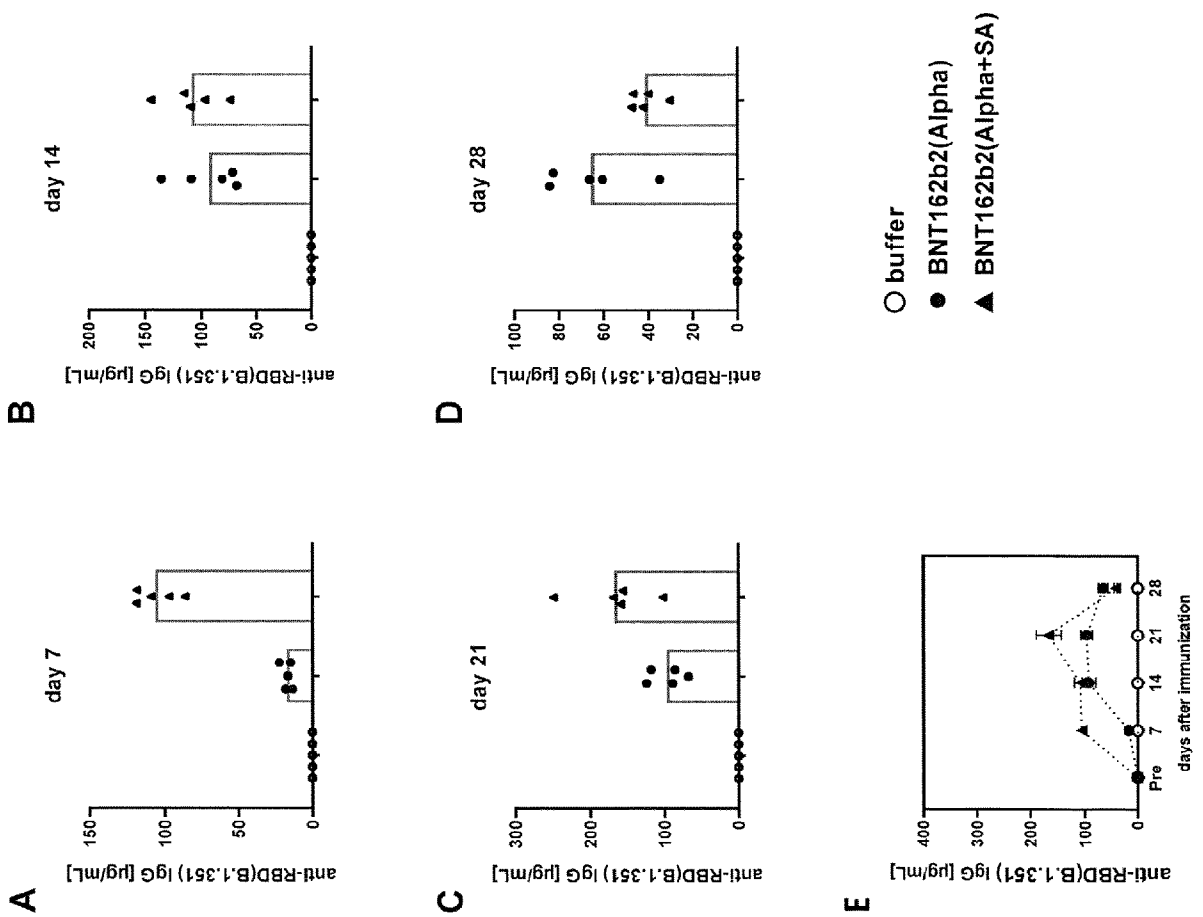

A
day 7
anti-RBD(B.1.351) IgG [µg/mL]
0
50
100
150
B
day 14
anti-RBD(B.1.351) IgG [µg/mL]
0
50
100
150
200
C
day 21
anti-RBD(B.1.351) IgG [µg/mL]
0
100
200
300
D
day 28
anti-RBD(B.1.351) IgG [µg/mL]
0
20
40
60
80
100
E
anti-RBD(B.1.351) IgG [µg/mL]
0
100
200
300
400
Pre
7
14
21
28
days after immunization
buffer
BNT162b2(Alpha)
BNT162b2(Alpha+SA)

Figure 10

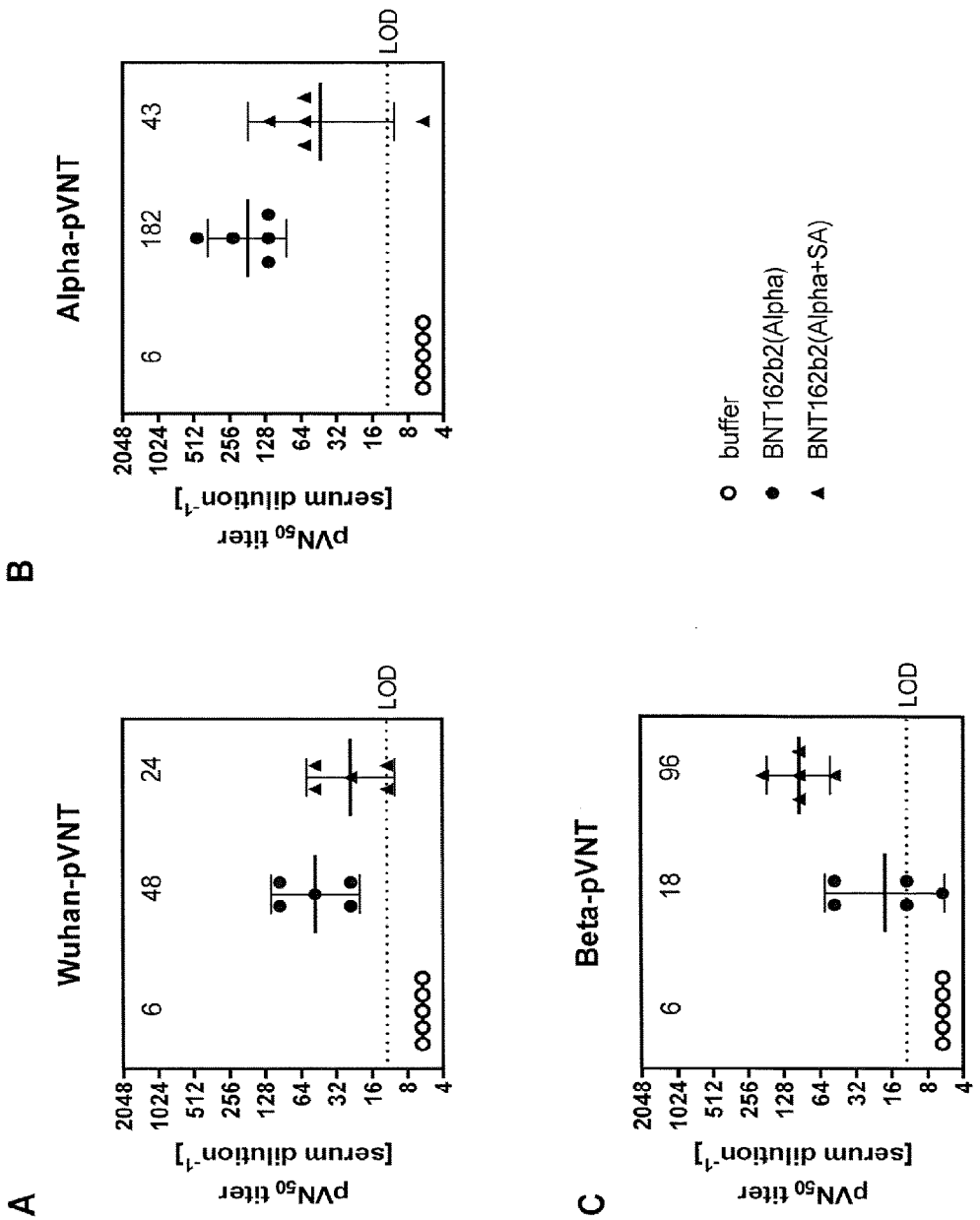

A
Wuhan-pVNT
B
Alpha-pVNT
C
Beta-pVNT
pVN50 titer
[serum dilution-1]
2048
1024
512
256
128
64
32
16
8
4
LOD
6
48
24
182
43
18
96
buffer
BNT162b2(Alpha)
BNT162b2(Alpha+SA)

Figure 11

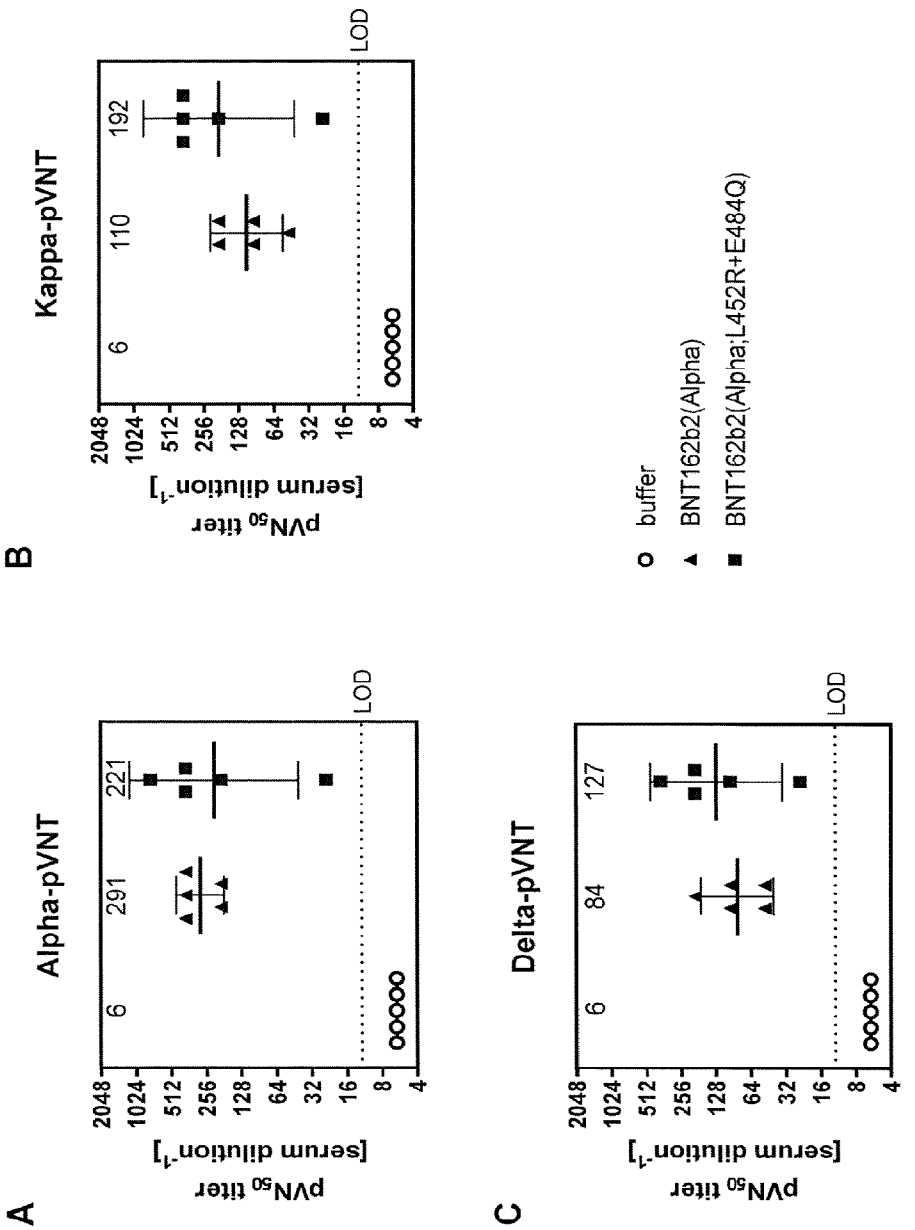
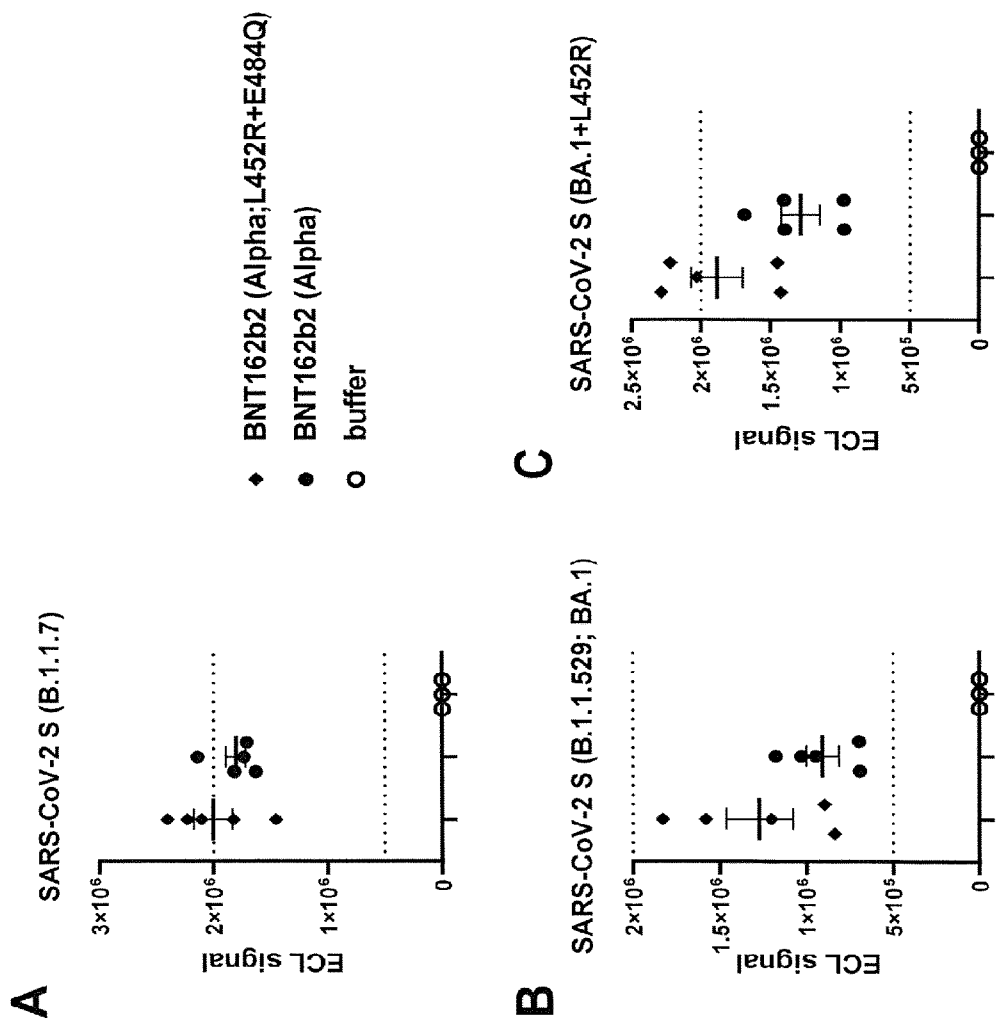

A
Alpha-pVNT
pVN50 titer
[serum dilution-1]
2048
1024
512
256
128
64
32
16
8
4
6
291
221
LOD
B
Kappa-pVNT
6
110
192
LOD
C
Delta-pVNT
6
84
127
LOD
buffer
BNT162b2(Alpha)
BNT162b2(Alpha;L452R+E484Q)

Figure 12


A
SARS-CoV-2 S (B.1.1.7)
ECL signal
3×10^6
2×10^6
1×10^6
0
B
SARS-CoV-2 S (B.1.1.529; BA.1)
ECL signal
2×10^6
1.5×10^6
1×10^6
5×10^5
0
C
SARS-CoV-2 S (BA.1+L452R)
ECL signal
2.5×10^6
2×10^6
1.5×10^6
1×10^6
5×10^5
0
BNT162b2 (Alpha;L452R+E484Q)
BNT162b2 (Alpha)
buffer